

```

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Landsberg erecta"
/db_xref="taxon:3702"
/chromosome="4"
/map="between apt and cer-9"
/tissue_type="flower; ovule"
1..1905
/gene="AINTEGUMENTA"
38..1705
/gene="AINTEGUMENTA"
/note="similar to APTAL2 protein encoded by Genbank
Accession Number U12546"
codon_start=1
/product="ANT"
/protein_id="AA017364.1"
/db_xref="GI:1244708"
/translation="MRSFCNDNNHNTNTNLLIGFSLSNMNRKGRGREGAIVSSST

```

US-10-024-632-2 (1-663) X ATU41339 (1-1905,

QY	1	MetLysArgIleAsnGluSerAsnAsnThrAspAspArgIleAsnAsnLysAsnTrpLeuGly	20
Db	38	ATCGAAGCTCTTTTGGGAGATGATGATATTAATCATGCAACACAGACTAATTGTGAGG	97
OY	21	PhaSerLeuSerProHis--MetLysMet	29
Db	98	TTCTCATTTGCTTCACAAATGCAATGCAAAATGGAGAGTAGAGAGGTAGAGAACTATTAC	15
OY	30	-----GluAlaThrSerAlaAlaThrValProThrThrPheTyMet	43
Db	158	TCATCTTCAACTCTTCAGCTGCAACTCTCTCTCTCTCTGTCACCTCAACTGTGTGT	217
OY	44	SerProSerGlnSerHisLeuSerAsnPhaGlyMetCysTyArgIleValGluAlaAsnGly	63
Db	218	GGT-----GACAAACACTGACCACTTGGTGGTTTGGTATGATGATCAACCCAAATGGA	266
OY	64	AsnPhaHisSerProLeuThrValMetProLeuLysSerAspGlySerLeuCysIleLeu	83
Db	269	GGAACTGATTTCCACATGCTGTATGAGCACTCAATGTGATGGTTCTTTTGCTTAATG	322
OY	84	GluAlaLeuLysArgSerGlnThrGlnValMetValProThrSerSerProLysLeuGln	103
Db	329	GAAGCCTCAACAGATCTTCTCACTCGAATGACACCATCAAGATTATCTCCAAAGTGGAG	388
OY	104	AspPheLeuGluGlyAlaThrMetGlyThrHisGlnGlyArgIleSerHisGluValGly--	122
Db	389	GATTTCTTTGGG-----ACCATATCAACAACAACAATGCAAAAGAGGCATG	433
OY	123	-----LeuSerLeuLysPheSerLeuTyArgThrAsnSerGlnAsnAlaGluAlaGlnProAsnArg	144
Db	437	GATCTTAAGCTGAGATGATGTTTATCTTCAACACCAACTAT-----GAGCCCAACACG	487
OY	142	AspLeuLeuSerGlnProPhe-----ArgGlnGlnGlyHisMetSerValGln	157
Db	488	ACTACAAACTTTCACAGAGTCTTTAGCTTCCTCCCTCAACCAAGAACCAT-----	533
OY	158	ThrHisProTyLysSerGlyLeuAlaCysHisGlyLeuTyArgAlaProLeuGlnGln	177
Db	536	-----GAGGA	544
OY	178	GluThrThrLysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGln	197
Db	542	GAAACT-----AGAAATTAACGGGAATATAC-----CCATGTTTATACAT	580
OY	198	GlyLeuLysAsnTrpValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGln	217

Db 581 GGA----- 583  
 Qy 218 GlnGlnMetAsnCysGlyMetGlyGlnGlnValArgGlnGlnValSerLeuGlySerValGly 237  
 Db 584 GGGCTTTTAAATGATGAGGATATGAGGAAATTTCAA----- 619  
 Qy 238 CysGlyGlnLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysVal 257  
 Db 620 -----CAGTCACTGAGCTTATCATGAGCCCGGCTGACATCTAGCTGCATC 667  
 Qy 258 ThrAlaProSer----- 261  
 Db 668 ACTGGCTCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 727  
 Qy 262 -----GlyThrAspSerValAlaVal 268  
 Db 728 CAGATCTGAGAGCTCTGAG 787  
 Qy 269 AspAlaLeuLeuArgGlyHis-----AlaLeuLeuGlnGlnGlnProValHis 285  
 Db 788 GCCAAG 847  
 Qy 286 ArgLeuSerLeuAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValThrArgHis 305  
 Db 848 AGAAATCTATCATCTTGTGACACAGACTTCTCATACCGAGGCTTACAGAGACT 907  
 Qy 306 ArgGlyThrGlyValArgGlyValHisLeuTyrAspAsnSerCysLeuLeuGlnGln 325  
 Db 908 AGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 967  
 Qy 326 ThrArgGlySerGlyValGlnValTyrLeuGlnGlyTyrAspMetGlnGlnLeuAlaArg 345  
 Db 968 AGTGAAG 1027  
 Qy 346 AlaTyrAspLeuAlaAlaLeuTyrTyrGlyProSerThrHisLeuAsnProSerLe 365  
 Db 1028 GCATATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1087  
 Qy 366 GlnAsnTyrGlnValGlnLeuGlnGlnMetLeuAsnMetSerArgGlnGlnTyrValAla 385  
 Db 1088 GAGAAATTTATCGAAG 1147  
 Qy 386 HisLeuArgArgLeuSerSerGlyPheSerArgGlyAlaSerLeuTyrArgGlyValThr 405  
 Db 1148 CATTTGAG 1207  
 Qy 406 ArgHisGlnHisGlyArgTyrGlnAlaArgGlyArgValAlaGlyValAsnLeuAsp 425  
 Db 1208 AGACATCACAG 1267  
 Qy 426 LeuTyrLeuGlyThrPheSerThrGlnGlnGlnAlaGlnAlaTyrAspValAlaAla 445  
 Db 1268 CTCACTGAG 1327  
 Qy 446 IleuAspPheArgGlyAlaAsnAlaValThrAsnPheAspIleuSerArgTyrAspValGln 465  
 Db 1328 ATTAAGTTCCGTGACAAATGCTGATCACTTGAATCAAGAGTACAGATGTTGAT 1387  
 Qy 466 ArgIleuMetAlaSerSerAsnLeuLeuAlaGlyGlnLeuAlaArgArgIleuAspAsn 485  
 Db 1388 CGTATCATGCTTATGATCAACCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1438  
 Qy 486 AspProArgAsnLeuAspIleuAspTyrAsnLeuSerValThrSerValAsnGln 505  
 Db 1439 -----AACACAGCATGTC-----GTCAAGAAATACT 1465  
 Qy 506 GlnThrValGlnValGlnAlaGlyAsnAsnAsnAsnGlnAsnAspSerGlnTyrIleuMet 525  
 Db 1466 GAA----- 1468  
 Qy 526 ValLeuPheAsnHisProSerGlnGlnGlnAlaGlnGlyAsnGlySerAspGlnLeu 545  
 Db 1469 -----GACCAAAAC 1477

Qy 546 IleuPheAsnGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565  
 Db 1478 GCTCTAAATGCT----- 1489  
 Qy 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlnSerSerIleGly 585  
 Db 1489 ----- 1489  
 Qy 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerArgGlnAlaSer 605  
 Db 1490 -----GTTGTGAAG 1522  
 Qy 606 ProGlnArgGlyProSerLeu-----LeuPheProMetProProMetGlnThrIle 623  
 Db 1523 ACTCCGAG 1582  
 Qy 624 IleVal-----AsnProIleGlyThrSerValThrSerThrLeuProSerProThrValGln 642  
 Db 1583 ATGTTGAGATCAAAATATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1642  
 Qy 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTyrPheAsp 662  
 Db 1643 CTTAAG-----ACGAGGCTCTTACTTGCCTCAGATGCCGTTTCCGCTGCTGAG 1699  
 Qy 663 Thr 663  
 Db 1700 TCT 1702  
 PRTIT 9



1	1328	37.9	555	2	S71365	ovule development
2	923.5	25.4	425	2	B66760	hypothetical prot
3	916.5	26.2	581	2	T15180	ovule development
4	890	25.4	485	2	T03638	hypothetical prot
5	879	25.1	516	2	P66549	hypothetical prot
6	831	22.7	566	2	T49988	ovule development
7	645	18.4	332	2	D86295	hypothetical prot
8	616	17.6	308	2	H68627	protein F20H17.12
9	534	17.2	485	2	S54116	hypothetical prot
10	517	1.8	205	2	T47591	antigenmueta-1-like
11	517	1.8	432	2	A85436	ABER1A2 protein
12	497.5	1.2	436	2	B84845	probable Ap2 domain
13	486.5	1.9	464	2	C84666	probable Ap2 domain
14	483	1.8	446	2	T03981	ABER1A2-like prot-
15	467.5	1.3	433	2	T01574	indeterminate split
16	286.5	8.2	247	2	T06721	hypothetical prot
17	275	7.9	236	2	T02577	probable Ap2 domain
18	161.5	4.6	262	2	E96747	hypothetical prot
19	160	4.6	314	2	B66482	protein F5U5.5 (11
20	158.5	4.5	245	2	B66168	hypothetical prot
21	158	4.5	292	2	T09030	hypothetical prot
22	157.5	4.5	199	2	B66288	Fall 31 protein -
23	156.5	4.5	248	2	T07728	transcription fact
24	155	4.4	358	2	D96579	hypothetical prot
25	152.5	4.4	336	2	E84594	Ap2 domain trans
26	151.5	4.3	268	2	T04787	hypothetical prot
27	151.5	4.3	334	2	T00896	hypothetical prot
28	150	4.3	256	2	H96837	unknown protein T
29	150	4.3	295	2	T00399	probable Ap2 domain

30	147.5	4.2	343	2	P05607	hypothetical protein
31	147.5	4.2	1475	2	P66399	protein F17L21.22
32	146	4.2	503	2	T03541	hypothetical protein
33	144	4.1	212	2	T04580	hypothetical protein
34	144	4.1	225	2	T02433	DNA binding protein
35	143.5	4.1	272	2	T05015	hypothetical protein
36	143.5	4.1	277	2	T03927	DNA binding protein
37	141.5	4.0	344	2	T51329	DNA binding protein
38	141	4.0	236	2	B84718	hypothetical protein
39	140.5	4.0	171	2	T00432	probable AP2 domain
40	140	4.0	283	2	C84828	AP2 domain transcr
41	139.5	4.0	222	2	T52019	eukaryotic response
42	139	4.0	218	2	T01076	transcription factor
43	139	4.0	352	2	T51330	DNA binding protein
44	137.5	3.9	204	2	B86410	protein F3M18.20
45	137.5	3.9	328	2	G84826	hypothetical protein

## ALIGNMENTS

RESULT 1  
571365  
ovule development protein aintequamenta - Arabidopsis thaliana  
N:Alternate names: A22 domain-containing protein; protein T2819.30  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 22-Oct-1999  
C:Accession: 571365; 106022  
R:Kluener, K.  
submitted to the EMBL Data Library, November 1995

A:Accession: S1365  
A:Molecule type: mRNA  
A:Residues: 1-555 <KU>  
A:Cross-references: EMBL:U0425; NID:G1209099; PID:G1209099  
R:Bayan, M.; Van Der Schueren, J.; Chang, Y.J.; Voet, M.;  
submitted to the Protein Sequence Database, March 1999

A; Cross-references: EMBL:AL035709; GSPDB:GN00062; ATSP:T28119.30  
A; Experimental source: cultivar Columbia; BAC clone T28119

A;Gene: ATSP:T28I19.30; ANTI

A; Introns: 111/1; 289/2; 317/1; 320/1; 349/3; 391/2; 417/1

Query Match	Score	DB 2	Length
37.9%	1328	2	555

Matches 314; Conservative 69; Mismatches 113; Indels 226; Gaps 22

1 MKRINESNTDGNHNHNLGFSLSPH-MKM-----EATSAALVPIITFM

Db 1 MKSFCNDNDNHSNTTNLGLFSLSSNMKMGGRGGREALYSSSISSAAISSSSVFPQLV

44 SPSQSHLSNFGMCYGVGENGFHSPLI VMPUNSDGSLIJEALNRSDIQMVVFISBFUW

61 G--DNISNFGVCIGSNPNUGLISHMSVMFLKSDGDELMEKLNKSDHSHNLLQDOOFNVE

[illegible]

Case	Age	Sex	Occupation	Duration of illness (years)	Site of lesion	Pathological changes	Remarks
1	45	M	Farmer	10	Brain	Chronic inflammation	
2	52	F	Housewife	15	Brain	Chronic inflammation	
3	60	M	Teacher	20	Brain	Chronic inflammation	
4	68	F	Retired	25	Brain	Chronic inflammation	
5	75	M	Farmer	30	Brain	Chronic inflammation	
6	82	F	Housewife	35	Brain	Chronic inflammation	
7	88	M	Farmer	40	Brain	Chronic inflammation	
8	92	F	Housewife	45	Brain	Chronic inflammation	
9	95	M	Farmer	50	Brain	Chronic inflammation	
10	98	F	Housewife	55	Brain	Chronic inflammation	

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

**במקור: רשמי**

```

Db      230 QQISEALVETSVGFETTTAAAKKKGGQDVVVVQKQIVARHSIDTGGQTSQVRGTR 289
Qy      305 HMTGRYEALWDNSCKEKGQTRKGRQVYLGSDYMEERAAAYDLAALKWGPSTHIFS 364
Db      290 HMTGRYEALWDNSCKEKGSRKGRQVYLGSDYMEERAAAYDLAALKWGPSTHIFS 349
Qy      365 IENYQVLEEMKMSQEVAAHLRRKSGFSRGASITGVTRHHQHGMRARIGRVANR 424
Db      350 AENYQKEIEDMKMTQEVAAHLRRKSGFSRGASITGVTRHHQHGMRARIGRVANR 409
Qy      425 DYLGTFTSTOEBAEAAYDVAALKFRGANAVTNFDISRVDVERIMASNLGELARRKD 484
Db      410 DYLGTFTSTOEBAEAAYDVAALKFRGTNAVTFNFDITRYVDRIIMSNTLLSGELARRN 467
Qy      485 NDRPRNDIDYKSVTVSVNNEETVQVQAQNNNNDESEKMYLPHHPGQQQANNGSDQ 544
Db      468 -----NNSTV-VNNT-----DQ 479
Qy      545 KIMNCGYNNNSAFSNALODLIGDSVSGQHMLDESSKIGTFNTSLYTSLSSPA 604
Db      480 TALNA-----VEGGSNREV 494
Qy      605 SPEKRPGL-LFPMPMETKIV-NPIGTSVTSMLPSPTVQMRPSPALSLPLPFASWT 661
Db      495 STPERLLSPFAIFALPQVQKQVFGSNMGNNSPMTSNPVALIK-TVALTLPMVFAAMA 553
Qy      662 DT 663
Db      554 DS 555

RESULT 2
B96750
hypothetical protein F28P22.24 [imported] - Arabidopsis thaliana
C/DSpecies: Arabidopsis thaliana (mouse-ear cross)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #ext_change 31-Mar-2001
C/Accession: B96750
R/Theologas, A.; Ecker, J.R.; Palm, C.U.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.X.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lures, J.S.; Maiti, R.; Marzilli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; NUID:21016719; PMID:11130712
A/Accession: B96750
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-425 <STO>
A/Cross-references: GB:A005173; NID:96648171; PIDN:AAF21171.1; GSPDB:GNC00141
C/Genetics:
A/Gene: F28P22.24
A/Map position: 1

Query Match      26.4%; Score 923.5; DB 2; Length 425;
Best Local Similarity 43.4%; Pred. No. 1.5e-57;
Matches 214; Conservative 60; Mismatches 108; Indels 111; Gaps 16;

```

```

Qy      192 MPQVBS-----LKNVV---APTRERSTQVLEQNMCM---GNERNGVSL 233
Db      110 DPNLSNDYGGFEVSVSVFSLMLEQTPAPLSSHYTEBACTNNIHFNSBEFGYNT 169
Qy      234 -GSVGGELQSLSLNSPGSSQSVTAPSGT---DSVAVDK-KRGAKLQKQPVHR 287
Db      170 NGSM-----LSLALSHGACDLINESNVSAVEEPVVKDEKRLVYXPQYKESVPR 222
Qy      288 SIFTFGRTQYRGVTRHRTGTYEALWDNSCKEKGQTRKGRQVYLGSDYMEERAAAY 347
Db      223 SVDSYGRFISQYGVTRHRTGTYEALWDNSCKEKGQTRKGRQVYLGSDYMEERAAAY 282
Qy      348 DLALKWGPSTHIFSIENYQVLEEMKMSQEVAAHLRRKSGFSRGASITGVTRH 407
Db      283 DLALKWGPSTHIFLNPPLSYKEIEELNNMNOEFAVMLRRNSGFSRGASITGVTRH 342
Qy      408 HOGRMQARIGRVANRKYDLYGTFTSTOEBAEAAYDVAALKFRGANAVTNFDISRVD 467
Db      343 HOGRMQARIGRVANRKYDLYGTFTSTOEBAEAAYDVAALKFRGANAVTNFDIRYDKI 402
Qy      468 MASSNLGELAR 480
Db      403 CSSTIVDSQAK 415

RESULT 3
T51580
coulc development protein ainegmenta-like protein - Arabidopsis thaliana
N/Alternate names: protein T10B6_90
C/DSpecies: Arabidopsis thaliana (mouse-ear cross)
C/Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #ext_change 18-Aug-2000
C/Accession: T51580
R/Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Aamizu, E.; Kotani, H.; Tabata, S.; M
submitted to the Protein Sequence Database, August 2000
A/Reference number: 225394
A/Accession: T51580
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-581 <SAT>
A/Cross-references: EMBL:AJ391142
A/Experimental source: cultivar Columbia; BAC clone T10B6
C/Genetics:
A/Map position: 5
A/Intons: 78/1; 216/2; 244/1; 273/3; 296/2; 315/2; 341/1
A/Note: T10B6_90

Query Match      26.2%; Score 916.5; DB 2; Length 581;
Best Local Similarity 35.2%; Pred. No. 7.5e-57;
Matches 255; Conservative 76; Mismatches 170; Indels 223; Gaps 25;

```

Db 202 ESPGRTSIRGVTRHRTGRYEALHMDNSCRREGOTRGRQ---GGYKEEKARAYDL 258

QY 350 AALKXGSPSTHINPSLENYOVOLIEEMKNSROEYVAHLRKSSGSPRASITRYRVTTHQ 409

Db 259 AALKXGTTTTPFSEYEKEEEMKNTROEYVAHLRKSSGSPRASITRYRVTTHQ 318

QY 410 HGRWQARIGRVAGNKDLYLGTFTSTOEAAEAYDVAALKFRGNAVTFNDSRYDVERIMA 469

Db 319 HGRWQARIGRVAGNKDLYLGTFTSTOEAAEAYDVAALKFRGNAVTFNDSRYDVERIMA 378

QY 470 SSULLAGELARRKNDPKNKIDYKSVYTVNNEETVOVQAGN---NNNENDEMKV 527

Db 379 SPSLPFGSSAKRLKD-----VNNPVPAMMTSNVSSANNVSGWQMTA 421

QY 528 FNNPS-----OQQANGSGDQKIMCNGYRNBAFSAALQDLIGDVSQGHMLDE 580

Db 422 FQHQGMDLTLQOQOERYVG---YNGGNLSTSTVCFK-----QEEEOQHFLNS 471

QY 581 SSKIGT--HRSNSSLVTSLSRSSRASPERSKPSLLPFMPMETKIVNPITSTVSNLPS 638

Db 472 PSMTVVDHSSITSDSVTCGVVSYGCGFPAL-----PVGTSV-NYDPF 517

QY 639 PTV-----QMRSP-----AISLSH-----LPVFAS 659

Db 518 TAAEIVANARNHYVAQHQQOQOIQSPGDFPVALSNHNSNMVFHGGEGGAGPFSV 577

QY 660 WTDI 663

Db 578 WNDI 581

## RESULT 4

T03638  
hypothetical protein - maize  
C/Species: Zea mays (maize)  
C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 20-Jun-2000  
C/Accession: T03638  
R/Daniel, T.J.  
submitted to the EMBL Data Library, January 1995  
A/Reference number: Z14985  
A/Accession: T03638  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-485 <DAN>  
A/Cross-references: EMBL:Z47554; NID:G1103747; PTDN:CAA87634.1  
A/Experimental source: strain Black Mexican Sweet

Query Match 25.4%; Score 890; DB 2; Length 485;  
Best Local Similarity 44.2%; Pred. No. 4.2e-55;  
Matches 200; Conservative 45; Mismatches 94; Indels 114; Gaps 10;

QY 187 DCSLMPQMTGLKXNVAPTRFSTHQVLEQNNCG--MGENANGS-----LG 234

Db 4 DMSAYPH-----HMLFSLSNNVHGLLEAFSSSGTPTGDEGAVESSPRVEDFLG 57

QY 235 SVG-----CGELQSLSLSMSPGSOSG-----CYTAPSGT 263

Db 58 GVCVGAAPRSRLQIRITSLCAASGSITARFLRYPAAGSGTTVGBPLSFTLAASST 117

QY 264 DSAVAVAKKGAHAKGQKQVPRKSIDTFCQRTSQYRGVTRHRTGRYEALHMDNSCKE 323

Db 118 DVAAASDDQS-----RSAETFGQRTSIRGVTRHRTGRYEALHMDNSCKE 165

QY 324 GQTRKRGQVYLGQYDMEERKAPAYDLAALKVGPSTHINPSLENYOVOLIEEMKNSROEY 383

Db 166 GQTRKRGQVYLGQYDMEERKAPAYDLAALKVGPSTHINPSLENYOVOLIEEMKNSROEY 225

QY 384 VAILRKSSGFSRGASTYRGVTRHRTGRYEALHMDNSCKE 443

Db 226 IAILRKSSGFSRGASTYRGVTRHRTGRYEALHMDNSCKE 285

QY 444 AALKFRGNAVTFNDSRYDVERIM-----ASSNLLAG 476

Db 286 AALKFRGNAVTFNDSRYDVERIM-----ASSNLLAG 476

QY 477 ELARKKNDPKNKI-----DYKSVTVNNEETVOVQAGN 515

Db 346 AAAAQATMPPEKQKYSWLLALHYQOQOQERQPPASAEYAYGGGVVDPMTGTSGSN 405

QY 516 NNENDS-EV---KXVLEFNHPQOQOANGSGD 543

Db 406 NNTGSGVWVGATGSAVVGQOQSSSKQNGVASN 438

## RESULT 5

P96549  
hypothetical protein FL1M5.6 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: P96549  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, N.R.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziani, R.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Ventner, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A6614; MUID:21016719; PMID:11130712  
A/Accession: P96549  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-516 <STO>  
A/Cross-references: GB:A6005173; NID:G4636931; PTDN:AAD30633.1; GSPDB:GN00141  
A/Genetics:  
A/Map position: 1

Query Match 25.1%; Score 879; DB 2; Length 516;  
Best Local Similarity 38.4%; Pred. No. 2.8e-54;  
Matches 217; Conservative 70; Mismatches 140; Indels 138; Gaps 17;

QY 100 PLEDFPLGATMGTEYSHRGLSLDSIYNSQNAARPRDLSPFPQOQGMSTVQTH 159

Db 15 PKYADFLGYSKSGDH-----HTDNLV 36

QY 160 PYSGLACHGLYQAPLEBETTKETHVSDCSLMPQMTGLKXNVAPTRFSTHQVLEEQ 219

Db 37 PY-----NDIHQ-----TNASDYVPQTNLPTVTCASN--AP-----NNYELQESA 77

QY 220 KMGCGNERNVSGVSGCGELSLSMSP-----GSQSCVTAFPSTGDSVAVDAKRGH 275

Db 78 HNL-----OSLTSMSGSTAAALAAEVAITYKASAEISADNSSTTTCGALV----- 124

QY 276 AVLGGQKQVPRKSIDTFCQRTSQYRGVTRHRTGRYEALHMDNSCKEOTRGRQVYLG 335

Db 125 -----EATPRRTLETFCQRTSIRGVTRHRTGRYEALHMDNSCRREGOTRGRQ--G 175

QY 336 GYDMEKAPAYDLAALKVGPSTHINPSLENYOVOLIEEMKNSROEYVAHLRKSSGFS 395

Db 176 GYDMEKAPAYDLAALKVGPSTHINPSLENYOVOLIEEMKNSROEYVAHLRKSSGFS 235

QY 396 RGASTYRGVTRHRTGRYEALHMDNSCKEOTRGRQVYLG 455

Db 236 RASAVYRGVTRHRTGRYEALHMDNSCKEOTRGRQVYLG 295

QY 456 NFDISRYDVERIMASNL-LAGELARRKNDPKNKIDYKSVTVNNEETVOVQAGN 515

Db 296 NFDISRYDVERIMASNL-LAGELARRKNDPKNKIDYKSVTVNNEETVOVQAGN 355

QY 501 SYNNEETVOVQ-----AGNNNE-----NDESEKMLFNHPQOQ 536

Db 356 SVASSRRLQOPPYPLQOPEHLHHHOPLLTQNNNDISQYHDSFYSYIOTLHQQOT 415



A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, K.R.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A:Reference number: A86141; MIMD:21016719; PMID:11130712

A:Accession: H96827  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-308 <STO>  
 A:Cross-references: GB:AF005173; NID:97715603; PIND:APF68121.1; GSPDB:GN00141

C:Genetics:  
 A:Gene: P20B17.12  
 A:Map position: 1

Query Match 17.6%; Score 616; DB 2; Length 308;  
 Best Local Similarity 50.0%; Pred. No. 5e-36;  
 Matches 132; Conservative 36; Mismatches 58; Indels 38; Gaps 7;

QY 254 SSCVTAPSGTDSVAVDAAKRGHAKLGGKOPVHRKSIDTFGORTQVGRVTRHRTGRYEA 313  
 DB 18 SOKTASASASIALTSKRRK-----RSPPRNAPL-----QRSSPIKGVTR-RTGRIYEA 65  
 QY 314 HMDNSCKEKGQTRKGRQVYLGVDMEKAAAYDLAALKYWGSPSTHINFSIENYOVQLE 373  
 DB 66 HMDKSNWMDTQTKKGRQ---GAYEEBAARAAYDLAALKYWGSDTLNPLPSYDEDVX 122  
 QY 374 EMKNSRQRYVAHLRKSQSGFSGASIRGYTRFHGHGRQWQARIGRVAGNKDLYLGTFTST 433  
 DB 123 EMEGSKERYISLRKSSGFSRGYSKRGVAHHHNRMAKRGVFGKLYLYLGYAT 182  
 QY 434 QEEAAEAYVAAIKRGANAATNPDISYDVERIYASSNLGLARKKDNDPR----- 468  
 DB 183 QEEAAIAYDIAIEYRGNAATNPDISYDVERIYASSNLGLARKKDNDPR----- 468  
 QY 489 ---NKDIDYKSVTVSNNEETVQ 509  
 DB 232 SREPSSDNK---SPKSEEVIE 251

RESULT 9  
 S54116  
 hypothetical protein - maize  
 C:Species: Zea mays (maize)  
 C:Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Sep-1997  
 C:Accession: S54116  
 R:Daniel, T.J.; Edwards, R.  
 A:Submitted to the EMBL Data Library, January 1995  
 A:Description: Complementation of a heat shock sensitive mutant of *Escherichia coli* defect

A:Reference number: S54116  
 A:Accession: S54116  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-485 <DAN>  
 A:Cross-references: EMBL:247554; NID:91103747; PID:91360703

Query Match 15.2%; Score 534; DB 2; Length 485;  
 Best Local Similarity 33.5%; Pred. No. 6.2e-30;  
 Matches 150; Conservative 48; Mismatches 146; Indels 104; Gaps 13;

QY 187 DCSLMPQMTBGLKXWVAATREFSTHQVLEQQNCG---MGNERNVGS-----LG 234  
 DB 4 DMSAYVPH-----HWLSPSLNNYHGLLEAFNSSGTPIGDDQGAVERSPRVEDPLG 57  
 QY 235 SVGC-GELOSLSLMSPGSQSCVTAPSGTDSVAVDAAKRGHAKLGGK-----QVTR-- 286  
 DB 58 GVCVAPAPRSRLQIR-----ITSLAASGCSITAFILHHYPAASGTTVGEPLSRFT 110  
 QY 287 ---KSIDTFGQRTSGYRGVTR-----FRMT-----GRYEAHLMDNSCKEKGQTRK 328  
 DB 111 LAAMSTDVAAVESDQASGPPRRSASAPSTVYASGTGGRDMRTCGNNSCRREGSKR 170  
 QY 329 GRQVYLGVDMEKAAAYDLAALKYWGSPSTHINFSIENYOVQLEEMKNSRQRYVAHLR 388  
 DB 171 GRQVYLGVDMEKAAAYDLAALKYWGSPSTHINFSIENYOVQLEEMKNSRQRYVAHLR 230

QY 389 RKSQSFSGASIRGYTRFHGHGRQWQARIGRVAGNKDLYLGTFTSTQEEAAEAYDVAIKF 448  
 DB 231 RKSQSFSGASIRGYTRFHGHGRQWQARIGRVAGNKDLYLGTFTSTQEEAAEAYDVAIKF 290

QY 449 RGAANAATNPDISYDVERIM-----ASNLGLAR 461  
 DB 291 PRACRHHQDMSRYDVESILSSDLPVGGAGSRAAAPPJDSLQPGSAAAMLAGAAAS 350

QY 482 KKNDPRKQDI-----DYNKSVTVSNNEETVQVQAGNNNNEND 520  
 DB 351 QATPFPSEKDYMSLLALHYQQQOQEQERQFPASAYEAAGSGGVVDFPMGTSSSGNNNTGS 410

QY 521 S-EW---KMYLFNHPSSQOQOQANNGSD 543  
 DB 411 GVMGATSGAVVGQODSSSKQNGYASN 438

RESULT 10  
 T47591  
 anti-gummaentra-like protein - Arabidopsis thaliana  
 N:Alternate names: protein T12E18.10  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C:Accession: T47591  
 R:Blocker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Queciet, F.; Salanoubat M. Mewes

A:Submitted to the Protein Sequence Database, March 2000  
 A:Reference number: Z24469  
 A:Accession: T47591  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-205 <BLO>  
 A:Cross-references: EMBL:AL132971  
 A:Experimental source: cultivar Columbia; BAC clone T12E18

C:Genetics:  
 A:Map position: 3  
 A:Introns: 71/2; 99/1, 128/3; 170/2  
 A>Note: T12E18.10

Query Match 14.8%; Score 517; DB 2; Length 205;  
 Best Local Similarity 55.1%; Pred. No. 2.7e-29;  
 Matches 109; Conservative 21; Mismatches 58; Indels 10; Gaps 4;

QY 241 LQSLSLMSPGSQ--SCVTAPSGTDSVAVDAAKRGHAK-----LGQOPVHRKSIDTFGCR 295  
 DB 5 LTTSTCSSSPSSSVSSSTTSSPIQSAAPPKRAKRAKSSPSGDKS--HNPTSPASTRK 62  
 QY 236 TSQYRGVTRHRTGRYEAHLMDNSCKEKGQTRKGRQVYLGVDMEKAAAYDLAALKY 355  
 DB 63 SSIYGVTRHRTGRYEAHLMDNSCKEKGQTRKGRQVYLGVDMEKAAAYDLAALKY 119  
 QY 356 GASTHINFSIENYOVQLEEMKNSRQRYVAHLRKSQSGFSGASIRGYTRFHGHGRQWQ 415  
 DB 120 GPDITLNPATVYKLEEMQRYKYEYLASLRQSSGFSRGYSKIRGVARRHHNGRMEA 179  
 QY 416 RIGRYAGNKDLYLGTFTST 433  
 DB 180 RIGRYAGNKDLYLGTFTST 197

# RESULT 11

A85436  
 APTAL2 protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
 C:Accession: A85436  
 R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spr  
 A:Title: Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.  
 A:Reference number: A85001; MIMD:20083488; PMID:10617198  
 A:Accession: A85436  
 A:Status: preliminary  
 A:Molecule type: DNA





QY 431 FSTCEAAEAYDVAIKFRGANAATNFDISRYDERIWAASNLAGEIARRKNDPRNK 490  
 DB 275 FDSEFEAAARADKAKALINNGREAVTNFEMSSYQNE-INSESN-----NS 317  
 QY 491 DIDYKSVVTSYVN-----EETVOVQAGNNNNENDSEMKVTLFNHP----SQQ 534  
 DB 318 EIDLNLIGISLSTGNPKONGRLFHPNSNTYERORG-VSLRIDNEYMKRPVTPLPYSSD 376  
 QY 535 QOANGSDOKIMNGCYNRNSAFSMALQ 562  
 DB 377 HRLYNNGA-----CPSYNNPAGGRATE 398

## RESULT 14

APETALA2-like protein Glossy15 - maize  
 C:Species: Zea mays (maize)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
 C:Accession: T03981  
 R:Moore, S.P.; Sisec, P.H.  
 Genes Dev. 10, 3018-3027, 1996  
 A:Title: Glossy15, an APETALA2-like gene from maize that regulates leaf epidermal cell  
 A:Reference number: Z15175; MUID:97115883; PMID:8957002  
 A:Accession: T03981  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-446 <MO>  
 A:Cross-references: EMBL:U41466; NID:91732030; PID:AAQ49567.1; PID:91732031  
 A:Experimental source: strain indred W64A; leaf  
 C:Genetics:  
 A:Gene: Glossy15

Query Match 13.8%; Score 483; DB 2; Length 446;  
 Best Local Similarity 31.9%; Pred. No. 2.2e-26;  
 Matches 119; Conservative 74; Mismatches 114; Indels 66; Gaps 10;

QY 250 PGSQSCVT---APSGTDSVAVDACKRGHAKLQ-----KQVHRKSIDTFGQRTSQY 299  
 DB 55 PDRRPAATVQOPFPPTTAACQATHEQCHVAGSABQWVSSASRSRRGRSSQY 114  
 QY 300 RGVTRHRTGRYEALHMDNSCKKEGQTRKRGCVYLGVDMEKAAAPADLAALKTWGST 359  
 DB 115 RGVTRHRTGRYEALHMDNSCKKEGQTRKRGCVYLGVDMEKAAAPADLAALKTWGST 359  
 QY 360 HINFESENQVOLEEKNNROEYVAHLRRKSSGFSRGASITRGVTRHHQHRQARIG 419  
 DB 165 DINFPLDDYDEKCKKDSKEBEFVTLRQAGAFVRSRRPGVTO-HKCGKMAKIGQ 223  
 QY 420 VAGNKDLYLGTSSTOEAAEAYDVAIKFRGANAATNFDISRYDERIWAASNLAGEI 479  
 DB 224 LMGKVVYVGLVTEFEAAQADYKAIKCYKEAVTNFDAQSYD-----267  
 QY 480 RKKKNDPRNKDIDYKSVVTSYVNNEETVOVQAGNNNNENDSEMKVTLFNHPSSQOQ 536  
 DB 268 -KELQGFWDGDELDELISGCASSDPTVAVEAFSPATSSSRKQPTMTLTGLPEEBST 326  
 QY 537 ANNGSDOKIMNGCYNRNSAFSMALQDILGIDSVSGGNNMDESSKIGTHFSNTSLVT 596  
 DB 327 GAG-----YFPAAGMFGRPADGHVVAAPRRHQWQOQOQ-QQH-----364  
 QY 597 SLSSREASPEKR 609  
 DB 365 ---AAPDAAPER 374

## RESULT 15

Indeterminate spikelet 1 - maize  
 C:Species: Zea mays (maize)  
 C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 29-Oct-1999  
 C:Accession: T01574  
 R:Chuck, G.; Meeley, R.B.; Hake, S.  
 Genes Dev. 12, 1145-1154, 1998

A:Title: The control of maize spikelet meristem fate by the APETALA2-like gene indeter  
 A:Reference number: Z14353; MUID:98221107; PMID:9553044  
 A:Accession: T01574  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-433 <CHU>  
 A:Cross-references: EMBL:AF048900; NID:92944039; PID:AAQ05206.1; PID:92944040  
 A:Experimental source: strain B73  
 C:Genetics:  
 A:Gene: Ids1  
 A:Map position: 1

Query Match 13.3%; Score 467.5; DB 2; Length 433;  
 Best Local Similarity 30.0%; Pred. No. 2.6e-25;  
 Matches 139; Conservative 68; Mismatches 164; Indels 93; Gaps 12;

QY 201 NMVAPTRBPTHQVLEQONCGMNERNGVSLGVCGLQSLSLMSPGSQSCVT-- 258  
 DB 6 NVASPADSGTSSSVL-----NSADGGFRFGL-LGSP--VDDDDCGEMAFGASTGFMTRQ 58  
 QY 259 -----APSGTDSVAVDACKRGHAKLQCKQVHRKSIDTFGQR 295  
 DB 59 LPSPPTPEPEPEEPAAPVPWQOPRADLGVQKPVAPAKVTRGP-----RSR 109  
 QY 296 TSQYRGVTRHRTGRYEALHMDNSCKKEGQTRKRGCVYLGVDMEKAAAPADLAALKTW 355  
 DB 110 SSQYRGVTRHRTGRYEMSHMD--C-----GKQVYLGSPDTAAARAYRAIKR 159  
 QY 356 GPSTHINFESENQVOLEEKNNROEYVAHLRRKSSGFSRGASITRGVTRHHQHRQAR 415  
 DB 160 GLDADINFLSDYEDDLKQRMNTKEFVHILRQSTGFARQSSKYRGVTL-HKCGRMA 218  
 QY 416 RIGRVAGNKDLYLGTSSTOEAAEAYDVAIKFRGANAATNFDISRYDERIWAASNL 475  
 DB 219 RMGOLGAKKITVGLDSEFEAAARADYRAALRNGREAVTNPESSYN-----AGDNNLR 273  
 QY 476 GELARRKNDPRNKDIDYKSVVTSYVNNEETVOVQAGNNNNENDSEMKVTLFNHPSSQO 535  
 DB 274 DTEFEALDDGDAIDLRLISQPNVOPKRDNTL---AGLQPTCSPSSSNTVMSQPMSSS 330  
 QY 536 QANGSDOKIMNGCYNRNSAFSMALQDILGIDSVSGGNNMDE-----580  
 DB 331 SPMP-----GTHQNPVVSFHQRLYSACHFPFHQVQEPVVERRPBLGQPP 380  
 QY 581 -----SSKIGTHFSNTSLVTSLSSSREASPEKRGPSTLFP 616  
 DB 381 SMWMAQSGPHVPLHHSAASSGFTAAAGANGWPLPSHPACPP 424

Search completed: March 9, 2004, 10:47:35  
 Job time: 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2004, 10:11:03 ; Search time 17 seconds  
(without alignments)  
2030.736 Million cell updates/sec

Title: US-10-024-632-2

Sequence: 1 MKRINENNTDGNHNNWIG.....RPSAISLSLVPASWTD 663

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	517	14.8	432	AP2_ARATH	P47927 arabidopsis
2	156.5	4.5	248	PT16_LYCES	O04682 lycopersico
3	155	4.4	5560	SPEN_DROME	Q08x83 drosophila
4	147	4.2	248	AP23_ARATH	P42736 arabidopsis
5	141.5	4.0	344	RAVI_ARATH	O9zwm9 arabidopsis
6	139.5	4.0	222	ERR4_ARATH	O80340 arabidopsis
7	134	3.8	243	ERR2_ARATH	O80338 arabidopsis
8	129.5	3.7	724	CC11_MOUSE	O9qgv9 mus musculu
9	129.5	3.7	727	CC11_MOUSE	O9xct6 equus caball
10	128.5	3.7	300	ERR5_ARATH	O80341 arabidopsis
11	127.5	3.6	1309	RAD3_YEAST	P14737 saccharomyc
12	124	3.5	2482	ATRX_PANTR	O7ygm4 pan troglod
13	122	3.5	1584	KYK1_DICDI	P18160 dictyosteli
14	121	3.5	725	CC11_PANTR	O8hxn7 pan troglod
15	121	3.5	2843	APC_HUMAN	P25054 homo sapien
16	120	3.4	726	CC11_HUMAN	O60563 homo sapien
17	119.5	3.4	161	PT15_LYCES	O04681 lycopersico
18	119.5	3.4	268	ERR1_ARATH	O80337 arabidopsis
19	118.5	3.4	918	YMB_CABEL	P34487 caenorhabdi
20	118.5	3.4	3075	LM1_HUMAN	P25391 homo sapien
21	118	3.4	1046	KUT4_YEAST	P47929 saccharomyc
22	115.5	3.3	1444	NCOS_HUMAN	O80339 arabidopsis
23	115	3.3	225	ERR3_ARATH	O90867 gallus gall
24	115	3.3	2492	HNFA_HUMAN	O9yev0 homo sapien
25	114.5	3.3	634	PCLO_HUMAN	O53591 streptococc
26	114.5	3.3	5147	HYSA_STRMA3	P18494 saccharomyc
27	113.5	3.2	984	GLN3_YEAST	P16462 actinobacill
28	113	3.2	730	LKTA_YEAST	P24339 schizosacch
29	113	3.2	1050	LKTA_YEAST	P24339 schizosacch
30	113	3.2	1085	GLI3_XENLA	O91660 xenopus lae
31	113	3.2	1569	GLI3_XENLA	O91660 xenopus lae
32	113	3.2	2492	ATRX_PONPY	O91660 xenopus lae
33	112.5	3.2	592	HN1B_XENLA	O91739 xenopus lae

## ALIGNMENTS

RESULT 1	ID	AP2_ARATH	STANDARD;	PRT;	432 AA.
AC	P47927	1996 (Rel. 33, Created)			
DT	01-FEB-1996	(Rel. 33, Last sequence update)			
DT	10-OCT-2003	(Rel. 42, Last annotation update)			
DE	Floral homeotic protein APTAL2.				
GN	AP2 OR ATG36920 OR AP22.49 OR C7A10.440.				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.				
OX	NCBI_TaxID=3702;				
RP	SEQUENCE FROM N.A. AND MUTANTS AP2-1 AND AP2-5.				
RC	SPRAIN-cv. Columbia; TISSUE=flower;				
RX	MEFLINE=95003709; PubMed=791989;				
RT	Yotoku K.D., den Boer B.G.W., van Montagu M., Okamura J.K.;				
RT	"Control of Arabidopsis flower and seed development by the homeotic				
RL	gene APTAL2.";				
RL	Plant Cell 6:1211-1225(1994).				
RP	SEQUENCE FROM N.A.				
RC	SPRAIN-cv. Columbia;				
RX	MEFLINE=20083486; PubMed=10617198;				
RA	Mayer K.F.X., Schueller C., Wambolt R., Murphy G., Voickaert G.,				
RA	Pol T., Duesterhoeft A., Stiekema W., Entian K.D., Terry N.,				
RA	Harris B., Ansoyge W., Brandt P., Grivall L.A., Rieger M.,				
RA	Weischelgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,				
RA	Kreis M., Deiseny M., Puidomech P., Watson M., Schmidheini T.,				
RA	Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,				
RA	Vos P., Hohenstiel U., Zimmermann W., Weller H., Ridley P.,				
RA	Langham S.-A., McCullagh B., Bilham L., Robben J., Vandebussche F.,				
RA	Van der Schueren J., Gymnopoulos B., Chang Y.-J., Vandebussche F.,				
RA	Braeken M., Welfens I., Voet M., Bastiaens I., Aert R., Defoor E.,				
RA	Weltzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,				
RA	Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,				
RA	Moolman P., Klein lanchorst R., Rose M., Haut J., Koester P.,				
RA	Bennet S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,				
RA	De Keyser A., Bysmaert C., Gielen J., Villarroel R., De Clercq R.,				
RA	Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray Allen S.,				
RA	Clark L., Doggett J., Hall S., Kay M., Leonard N., McIlroy K., Mayes R.,				
RA	Peterson A., Rajendram M.A., Lyne M., Benes V., Rechmann S.,				
RA	Borkova D., Blocker H., Scharfe M., Grimm M., Loehart T.-H.,				
RA	Dose S., de Haan M., Marse A.C., Schaefer M., Meller-Auer S.,				
RA	Gabel C., Fuchs M., Fattmann B., Granderath K., Dauner D., Herzl A.,				
RA	Neumann O., Argilou A., Vitale D., Ligotti R., Piravandi B.,				
RA	Wassenaar O., Argilou F., Clabaud G., Mwendile A., Felder R.,				
RA	Schnabl S., Hiller R., Schmidt W., Leclatry A., Aubourg S.,				
RA	Chedori F., Cooke R., Berger C., Monfort A., Casacuberta E.,				
RA	Gibbons T., Weber N., Vandenbol M., Bagnues M., Terol J., Torres A.,				
RA	Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,				
RA	Heijnen L., Schwarz S., Scholler P., Heber S., Franck P., Bielke C.,				
RA	Fishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,				
RA	Zaccaria P., Bevan M., Wilson R.K., de la Bastide V., Habermann K.,				

34	112	3.2	1211	1	BUN2_DROME	Q24523 drosophila
35	112	3.2	1244	1	MIX1_CABEL	Q09591 caenorhabdi
36	111.5	3.2	676	1	SP1_COTJA	P23499 colutrix co
37	111.5	3.2	3060	1	BPEB_HUMAN	O8w4x8 homo sapien
38	110.5	3.2	667	1	CY11_BOVIN	P35662 bos taurus
39	110.5	3.2	970	1	PSU1_YEAST	P53550 saccharomyc
40	110.5	3.2	1963	1	TFD2_MOUSE	O9epq8 mus musculu
41	110	3.1	1305	1	FTSK_YEPE	O8z9c7 yersinia pe
42	109.5	3.1	774	1	YMAE_YEAST	Q04233 saccharomyc
43	109	3.1	1447	1	BUD4_YEAST	P47136 saccharomyc
44	108.5	3.1	967	1	KINH_LOLPE	P21613 loligo peal
45	108.5	3.1	1211	1	SBCC_PSEAE	Q9hwb8 pseudomonas



RA Parnell L., Dethia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,  
RA Stoeckling T., Kalicki J., Graves T., Harmon G., Edwards J.,  
RA Lattrell P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
RA Kramer J., Fulton L., Mardis E., Dante N., Geisel C., Layman D.,  
RA Nelson J., Spieth J., Ryan E., Andrews S., Gelpi C., Layman D.,  
RA Du H., Ali J., Bergshoeff A., Jones K., Dione K., Cotton K., Joshi C.,  
RA Antocioni B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,  
RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Matero A., Shah R.,  
RA Swaby I., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,  
RA Grant S., Shohdy N., Hasegawa A., Hameed A., Lodi M., Johnson A.,  
RA Chen E., Marra M.A., Marienissen R., McCombe W.R.,  
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
RT thaliana.";  
RL Nature 402:769-777(1999).  
RN [3]  
RP FUNCTION.  
RX MEDLINE=91256319; PubMed=1675158;  
RA Drews G.N., Bowman J.L., Meyerowitz E.M.;  
RT "Negative regulation of the Arabidopsis homeotic gene AGAMOUS by the  
RT AP2/LA2 product.";  
RL Cell 65:991-1002(1991).  
CC -1- FUNCTION: Probable transcriptional activator that promotes early  
CC floral meristem identity. Is required subsequently for the  
CC transition of an inflorescence meristem into a floral meristem.  
CC Plays a central role in the specification of floral identity,  
CC particularly for the normal development of sepals and petals in the  
CC wild-type flower. Acts as a class cadstral protein by repressing  
CC the C-class floral homeotic gene AGAMOUS in association with  
CC others repressors like LEUNIG and SEUSS. It is also required  
CC during seed development.  
CC -1- SUBUNIT: May form homodimer.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- TISSUE SPECIFICITY: Sepals, petals, stamens, carpels, developing  
CC ovules, inflorescence stem, leaf, and stem.  
CC -1- DEVELOPMENTAL STAGE: It is detectable at low levels throughout the  
CC shoot apex, and at enhanced levels in the inflorescence meristem,  
CC young floral buds, and throughout the early stages of flower  
CC development and organogenesis. During floral organ differentiation  
CC it becomes spatially restricted to specific organ, tissue, and  
CC cell types within the flower.  
CC -1- INDUCTION: Negatively regulated by the C class floral  
CC homeotic protein AGAMOUS in stamens and carpels.  
CC -1- MISCELLANEOUS: Mutations in the AP2/LA2 gene result in the  
CC ectopic expression of AGAMOUS, leading to the replacement of  
CC sepals by carpels and stamens and of petals by stamens.  
CC -1- SIMILARITY: Contains 2 AP2/ERF domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: U12546; AAC13770.1; -;  
CC EMBL: Z99707; CAB16765.1; -;  
CC EMBL: AL161590; CAB80358.1; -;  
CC PIR: A85436; A85436.  
CC TRANSFAC: T01774; -;  
CC InterPro: IPR001471; TF\_ERF.  
CC Pfam: PF00847; AP2-domain.2.  
CC PRINTS: PR00367; ETRFSELEMENT.  
CC ProDom: PD001423; TF\_ERF\_1.  
CC SMART: SMO0380; AP2\_2.  
CC Flowering: Transcription regulation; Activator; Developmental protein;  
CC Nuclear protein; DNA-binding; Repeat.  
CC DOMAIN 14 50 ASP/GLU/SER-RICH (ACIDIC; POTENTIAL  
CC INVOLVEMENT WITH TRANSCRIPTION).  
CC FT DOMAIN 119 128 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
CC FT DNA\_BIND 128 192 AP2/ERF 1.

FT DNA\_BIND 220 286 AP2/ERF 2.  
FT MOTAGEN 159 159 G->E: TEMPERATURE SENSITIVE; AP2-5.  
FT MOTAGEN 251 251 G->S: TEMPERATURE SENSITIVE; AP2-1.  
FT MOTAGEN 420 420 O->E: TEMPERATURE SENSITIVE; AP2-5.  
SQ SEQUENCE 432 AA; 47833 MW; 376A355291EFPB42 CRC64;  
Query Match 14.8%; Score 517; DB 1; Length 432;  
Best Local Similarity 31.3%; Pred. No. 1.4e-28;  
Matches 163; Conservative 66; Mismatches 150; Indels 142; Gaps 19;  
171 YQAPLEETKETH-----VSDCSLMPQTEGLK---NWAPFRESTHQVLE 217  
9 HQQRESESEFECYSSPSKRVGSFNSSSSAVVIDGSDDEINRFRPNPLVTHQFPBE 68  
218 QQNNCGNERNRNVSLGSCGLQSLSLMSNGS-----QSSCVTAPSGTDSVAVD 269  
69 MDSNGG-----GVASG-----FPRAMFGVKFCQSLATGSSAGKATNVA 108  
270 AKKRGHAKGQKQPVHRKSIDTFGORTSQYRGVTRRWGTGRYBAHLMDNSCKEQTGRK 329  
109 A-----AVPEAPQL-KKSRGRGPRSSQYRGVTFYFRTRGMRSHIMD--C-----G 152  
330 RQYVLGYDMEERKARAYDLAAKRWGPSTHINPSTENVOLEMKMSRCQYVAHLR 389  
153 KQYVLGFDFDPAHAARAAYDPAIKFQGVENDIPNDIDYDDDKQMTNLTKESFVHLR 212  
390 KSGFSRGAFTYGVTRHHQGRWQARIGVAGNKDLYLGFSTCEAAEAAYDVAIKR 449  
213 QSTFPFGSSKRYGVL-HGCRWEANMGFLKRYVYGLFTEVEAARAAYDKAIKCN 271  
450 GANAATVPDISRDVEEYMASSNULAGELARKKNDPRKQIDYKNSVYTSVNNBETV 509  
272 GKDAVTFPDSIYDEELNAESSG-----NPTTPQDND----- 305  
510 VQAGNNNENDESEWKVLFNHPSSQQQANGSGDQKIMNGYRNSAFNALDILGIDS 569  
306 LSLGNSANS-----HKSQDMLRNQOQD-----SLHSNEVGLGQ 343  
570 VGSQGH-----NMLDESKTGT--HF-----NNSLVTSLSGSRASPEK 608  
344 TGMNHTPNSNHCPFGSSNIGSGGFSLPFAENHREFDGRASNTQVLTNAASGFSPPH 403  
609 RGPSLTFPMPMETKYNPIGTSVTSMLEPSTVQ---MRPS 646  
404 -----HNOITNSISTPROMVLQTNQPGPPLMRPS 432  
RESULT 2  
ID PT16 LYCES STANDARD, FRT, 248 AA.  
AC 004682;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Pathogenesis-related genes transcriptional activator PT16.  
GN P16.  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97357308; PubMed=92114637;  
RA Zhou J., Tang X., Martin G.B.;  
RT "The Pto kinase conferring resistance to tomato bacterial speck  
RT disease interacts with proteins that bind a cis-element of  
RT pathogenesis-related genes.";  
RT EMBO J. 16:3207-3218(1997).  
RL -1- FUNCTION: Transcription factor that binds to the GCC-box  
CC pathogenesis-related promoter element. Activates plants defense  
CC genes.  
CC -1- SUBUNIT: Interacts with the Pto kinase.

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: Contains 1 AP2/ERF domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U89257; AAC49741.1; -  
 CC PIR: T07728; T07728.  
 CC HSSP: O80337; 29GC.  
 CC TRANSFAC: T04717; -  
 CC InterPro: IPR01471; TF\_ERF.  
 CC Pfam: PF00847; AP2-domain; 1.  
 CC PRINTS: PR00367; ETRHSPLENT.  
 CC ProDom: PD001423; TF\_ERF; 1.  
 CC SMART: SM00380; AP2; 1.  
 CC Transcription regulation; DNA-binding; Activator; Nuclear protein.  
 KW DNA\_BIND 95 159 AP2/ERF  
 FT SEQUENCE 248 AA; 27913 MW; 115BD44DEB1B992 CRC64;  
 SC  
 Query Match 4.5%; Score 156.5; DB 1; Length 248;  
 Best Local Similarity 29.0%; Pred. No. 0.00072;  
 Matches 47; Conservative 16; Mismatches 54; Indels 45; Gaps 4;  
 QY 359 THINSIENYQVQLEMKMSQEQEYVALRKSSFSGASTYGVTRHHQGMORIG 418  
 DB 67 TEINIMPSTKSGDKRRKVSPTS--RRKK-----FGV-RQPPWGMALAIR 114  
 QY 419 RVAGNKDYLGTFTSTOEAAEAYDAIKFGANAVTFDSRYVERINASSNLLAGEL 478  
 DB 115 DTRGRKRWLGTTPDEAAVYDGAAYKLNQPDVNTFPVST----- 157  
 QY 479 ARRKQNDPRNDIDYNSVTSVNNETVQVQAGNNEND 520  
 DB 158 -----TAEVTVTETETESVADGDKSEND 183  
 RESULT 3  
 SPEN DROME STANDARD; PRT: 5560 AA.  
 ID SPEN\_DROME Q9NHN1; Q9NHN1; Q9VPL2;  
 AC Q8SXB3; Q9NHN1; Q9NHN1; Q9VPL2;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Split ends protein.  
 GN SPEN OR CG18497.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RX MEDLINE=20025996; PubMed=10556062;  
 RA Welleite E.L., Harding K.W., Mace K.A., Ronehaugen M.R., Wang F.Y.,  
 RA McGinnis W.;  
 RT 'Spn encodes an RNP motif protein that interacts with Hox pathways  
 RT to repress the development of head-like sclerites in the Drosophila  
 RT trunk.';  
 RL Development 126:5373-5385(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 4).  
 RC TISSUE=Embryo;  
 RX MEDLINE=2017049; PubMed=10655223;  
 RA Rebay T., Chen F., Hsiao F., Kojodziej P.A., Kuang B.H., Laverly T.,  
 RA Sun C., Voas M., Williams A., Rubin G.M.;

RT "A genetic screen for novel components of the Ras/mitogen-activated  
 RT protein kinase signaling pathway that interact with the Yan gene of  
 RT Drosophila identifies split ends, a new RNA recognition motif-  
 RT containing protein.";  
 RL Genetics 154:695-712(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), SUBCELLULAR LOCATION, AND TISSUE  
 RP SPECIFICITY.  
 RC TISSUE=Embryo;  
 RX MEDLINE=20171275; PubMed=10704397;  
 RA Kuang B.H., Wu S.C.-Y., Shin Y.-A., Luo L., Kojodziej P.A.;  
 RT "Split ends encodes large nuclear proteins that regulate neuronal  
 RT cell fate and axon extension in the Drosophila embryo.";  
 RL Development 127:1517-1529(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams W.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abri'l J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,  
 RA Ballou R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Bereson K.Y., Bencs P.V., Berman B.P., Bhattacharya D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,  
 RA Fjosek A., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Foster C., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heilmann T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,  
 RA Lascko P., Lei Y., Levitsky A.A., Li U.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Malishina N.V., Mobarry C., Morris J., Mostrel A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson X.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrinskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach U.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [5]  
 RP REVISIONS, AND ALTERNATIVE SPLICING.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bellen Court B.R., Ceiniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harpale N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.F.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [6]  
 RP SEQUENCE OF 424-2002 FROM N.A.  
 RC STRAIN=Berkley; TISSUE=Embryo;  
 RX MEDLINE=22426066; PubMed=12537569;

RA Stapleton R.A., Carlson J.W., Brokstein P., Yu C., Champe M.,  
 RA George R.A., Garin H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,  
 RA Rubin G.M., Ceiniker S.E.,  
 RA "A Drosophila full-length cDNA resource."  
 RT Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8 (2002).  
 RP [7]  
 RP FUNCTION.  
 RX MEDLINE=20253107; PubMed=10790398;  
 RA Lane M.E., Blend M., Heidmann D., Herr A., Marzock S., Herzig A.,  
 RA Lehner C.F.,  
 RT "A screen for modifiers of cyclin E function in Drosophila  
 RT melanogaster identifies Cdk2 mutations, revealing the insignificance  
 RT of putative phosphorylation sites in Cdk2."  
 RL Genetics 155:233-244(2000).  
 RN [8]  
 RP FUNCTION ON EGF RECEPTOR PATHWAY.  
 RX MEDLINE=20414403; PubMed=10959845;  
 RA Chen F., Rebay I.,  
 RT "Split ends, a new component of the Drosophila EGF receptor pathway,  
 RT regulates development of midline glial cells."  
 RL Curr. Biol. 10:943-946(2000).  
 RN [9]  
 RP FUNCTION ON Wg PATHWAY.  
 RX MEDLINE=22668876; PubMed=12783785;  
 RA Lin H.V., Dorquez D.B., Cho S., Chen F., Rebay I., Cadigan K.M.,  
 RT "Split ends is a tissue/promoter specific regulator of WINGLESS  
 RT signaling."  
 RL Development 130:3125-3135(2003).  
 CC -1- FUNCTION: Probable corepressor protein, which regulates different  
 CC key pathways such as the EGF receptor and Wg pathways. Involved in  
 CC neuronal cell fate, survival and axon guidance, cell cycle  
 CC regulation and repression of head identity in the embryonic trunk.  
 CC May act with the Hox gene Deformed and the EGF receptor signaling  
 CC pathway. Positive regulator of the Wg pathway in larval tissues  
 CC but not in embryonic tissues. May act as a transcriptional  
 CC corepressor protein, which repress transcription via the  
 CC recruitment of large complexes containing histone deacetylase  
 CC proteins.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative promoter;  
 CC Comment=2 isoforms, 1 (shown here) and 2, are produced by use of  
 CC alternative promoters;  
 CC Event=Alternative splicing; Named isoforms=4;  
 CC Name=1;  
 CC IsoId=Q8SX83-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q8SX83-2; Sequence=VSP\_008565; VSP\_008566;  
 CC Note=No experimental confirmation available;  
 CC Name=3; Synonyms=Spenu;  
 CC IsoId=Q8SX83-3; Sequence=VSP\_008567;  
 CC Note=Produced by alternative splicing of isoform 1;  
 CC Name=4; Synonyms=Spens;  
 CC IsoId=Q8SX83-4; Sequence=VSP\_008565; VSP\_008566; VSP\_008567;  
 CC Note=Produced by alternative splicing of isoform 2;  
 CC -1- TISSUE SPECIFICITY: Ubiquitous. Expressed prior to cellularization  
 CC in stage 3 embryos, and in blastoderm cells, including pole cells.  
 CC Expressed throughout the rest of embryogenesis. Later, it is  
 CC expressed at higher level in epidermal cells and CNS.  
 CC -1- DEVELOPMENTAL STAGE: Isoform 3 is expressed both maternally and  
 CC zygotically.  
 CC -1- SIMILARITY: Belongs to the Spen family.  
 CC -1- SIMILARITY: Contains 3 RNA recognition motif (RNM) domains.  
 CC -1- SIMILARITY: Contains 1 SPOC domain.  
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-7 is the initiator.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF188205; AAF1318.1; -  
 DR EMBL; AF184612; AAF26299.1; -  
 DR EMBL; AF221715; AAF34661.1; ALT\_INIT.  
 DR EMBL; AE003590; AAF51534.2; -  
 DR EMBL; AE003590; AAF51535.2; -  
 DR EMBL; AE003590; AAF10511.1; -  
 DR EMBL; AY094789; AAM1141.1; ALT\_SEQ.  
 DR HSSP; P09651; 1HA1.  
 DR FlyBase; FBgn0016977; spen.  
 DR GO; GO:0007411; Paxon guidance; IMP.  
 DR GO; GO:0008347; P-glia cell migration; IMP.  
 DR InterPro; IPR00504; RNA\_rec\_mob.  
 DR Pfam; PF00076; Rnm; 3.  
 DR SMART; SM00360; RRM; 3.  
 DR PROSITE; PS50102; RRM; 3.  
 DR PROSITE; PS00030; RRM RNP 1; FALSE\_NEG.  
 DR PROSITE; PS50917; SPOC; 1.  
 DR Transcription regulation; Repressor; Developmental protein,  
 KM Nuclear protein; Repeat; RNA-binding; Coiled coil;  
 KM Alternative promoter usage; Alternative splicing.  
 FT DOMAIN 554 632 RNA-BINDING (RRM) 1.  
 FT 554 730 RNA-BINDING (RRM) 2.  
 FT DOMAIN 656 730 RNA-BINDING (RRM) 3.  
 FT DOMAIN 734 806 RNA-BINDING (RRM) 3.  
 Query Match 4.4%; Score 155; DB 1; Length 5560;  
 Best Local Similarity 19.7%; Pred. No. 0.081;  
 Matches 144; Conservative 85; Mismatches 266; Indels 236; Gaps 30;  
 7 SNTDGGNNNNWGLSISPMKMEATSATVPTFTYMSBQSHTLSFNGCYGNGNPH 66  
 358 TNNFSNSNNTTTPPTPIAGAG--SVLGSGAAGVCSNGS----- 404  
 67 SPLVMLKSDGSLCLEAKSQOTVMVPTSSPKLEDFGATMGHEGHERGLSD 126  
 405 -----TSGDLNAAVLAADVNGVTHPIRIRNLHGRT--TSSRSRSRSPSV 455  
 127 SIYVSNAAEAPNRDLISQPRQGHMSVQTHPVYSGLAHGLYQAPLEETTKETHVS 186  
 456 SSSSSSSSSSHSSSHSSSHSSSPVQSG----- 481  
 187 DCSSLMQMTGELK---NVAAPTRERSTHQVLEQGMNGENENGVSLSVGGEELQ 242  
 482 NCA-----MAEGSSRTVNSVTVTSNNSPGTAIVTSAGVG-----GGS-S 525  
 243 SLISMSPGSQSGCVTA-----PSGTDVAVDARKKHAK 277  
 526 SSSSSSSSSSGSGLTANPVVHSEDNRPLAIRVNLPARSSDLSKDGLFHEYKKGKVT 585  
 278 LGQKQPVHRKSIDTFQGTQYKGYTHRWGTGYEALMNSC-KREGQTRKGRV--- 332  
 586 W-----VKVVGNSERYALV-----CFKKPDVEKALEVSHDK 618  
 333 -YLG-----GYDNEEAARAYDLAALKYGPSTHINFSTENYOVULEMKMSRQZ 382  
 619 HFPGCKEVEYQGYDVEDNEFRYEALDEHPKSRITLF-IGNLE-----KDIAGE 671  
 383 VVAHLRR-----KSG-----FSGASITYGVT---HHQGMQARIGRA 421  
 672 LRSHEAFGEIIEIDIKQGINAVAFQYSDIVSVVAMKMGDEHLSNRILKGFQSK 731  
 422 GNKLYLGTFSTOEAAEAVDAIKFRGNAAVNPFISR-----YD-VERIMASSN 472  
 732 PTNCWMD--GVDEKVSSEF--LOSQTRRGATYKVSIDNRQLALVLYQVOQAQAVK 787  
 473 LLAELARRKNDPRNKDIDYKSVTVSVNNEVTVQAGNNNENDSEMKVLPFHS 532  
 788 DMRGTLIRKK-----LQVDPAARECD-----AFYDQ 816  
 533 QQQQANGSPQKIMNCGYNSAFSMALODLIGDVGSGQHMLDESSKIGT----- 586  
 817 EKQQQSSGSNPRF---SRYSSASS--LQSRASFSRHHQNNNSNDCCSPINPGASS 871

Cy 587 HFSNTSLVTSLSRSSRSPKGRSLLPMPMETKIVN-----PIGTSVSWLP 637  
 Db 872 GISSASLILNG-SISINISINIGNACSNAMPASLAAVSNVNASGTVFASSTSPGVS 930  
 Cy 638 SPTVQMPSPA 648  
 Db 931 SSSSLSPMSPA 941  
 RESULT 4  
 AP23\_ARATH STANDARD; PRT; 248 AA.  
 AC P47367.023105; 32, Created)  
 DT 01-NOV-1995 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE AP2 domain transcription factor RAP2.3 (Related to AP2 protein 3)  
 DE (Cadmium-induced protein AS30).  
 GN ATG16770 OR MG6.1 OR MG6.24.  
 OS Arabidopsis thaliana (mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Choi S., Baek E., Lee S.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=9738147; PubMed=9192694;  
 RA Okumuro J.K., Caster B., Villarreal R., van Montagu M., Jofuku K.D.;  
 RT "The AP2 domain of APT2A2 defines a large new family of DNA binding  
 RT proteins in Arabidopsis".  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:7076-7081(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA MEDLINE=20277480; PubMed=10819329;  
 RA Sato S., Nakamura Y., Kaneo T., Katoh T., Asamizu E., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
 RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC  
 RT clones".  
 RL DNA Res. 7:131-135(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA MEDLINE=22954850; PubMed=14593172;  
 RA Yanada K., Lim J., Dale J.M., Chen H., Shih N., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Piam P.K., Cheuk R.F.,  
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Miranda M., Quach H.L., Trapp C.S., Deng C.H., Lee J.N., Toriumi M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 RA Aizawa T., Bando J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Chiu A., Goldsmith A.D., Gurjani M., Hansen P.,  
 RA Hayashizaki Y., Johnson-Hopson C., Hsuang V.W., Iida K., Karnes M.,  
 RA Khan S., Koeseema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kamita A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
 RA Saito M., Tamse R., Vaysberg M., Wallender E.K., Kong C., Yamamura Y.,  
 RA Yuan S., Yoshizaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 RT genome".  
 RL Science 302:842-846(2003).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Brover V., Truham M., Alexandrov N., Lu Y.-P., Flavell R.,  
 RA Feldmann K.A.;  
 RT "Full-length cDNA from Arabidopsis thaliana".  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- INDUCTION: By cadmium.  
 CC -1- SIMILARITY: Contains 1 AP2/ERF domain.

CC -1- CAUTION: Ref.1 sequence is most probably a cloning artifact and  
 CC differs from that shown due to an insertion into the sequence.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation at  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 DR EMBL; Z37504; CAA5734.1; ALT\_SEQ.  
 DR EMBL; AF003096; AAC49769.1; -  
 DR EMBL; AB022217; BAB02769.1; -  
 DR EMBL; AY059917; AAL24399.1; -  
 DR EMBL; AY035100; AAK59605.1; -  
 DR EMBL; AY142562; AAN13131.1; -  
 DR EMBL; AY087489; AAM65031.1; -  
 DR HSSP; O80337; 2GCC.  
 DR TRANSFAC; T02658; -  
 DR InterPro; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 1.  
 DR PRINTS; PR00367; ETRSPLEMENT.  
 DR PRODOM; PD001423; TF\_ERF; 1.  
 DR SMART; SM00380; AP2; 1.  
 KW Transcription regulation; DNA-binding; Nuclear protein; Cadmium.  
 FT DNA BIND 76 141  
 FT DOMAIN 140 143  
 FT POLY-PRO. 140 143  
 SQ SEQUENCE 248 AA; 27758 MW; 317D50CC514782F1 CRC64;  
 Query Match 4.2%; Score 147; DB 1; Length 248;  
 Best local similarity 25.1%; Pred. No. 0.0033;  
 Matches 67; Conservative 34; Mismatches 100; Indels 66; Gaps 10;  
 Cy 328 KGRQVVLGGYDMEERAAAYDLAAKMG-----PSTHINPSIENVYQLEMKMS 379  
 Db 18 KGRKL-----TAEELMSLSDASADDFWGFSTSKLHPTNVN-----VKEARVKEQ 65  
 Cy 380 ROEYVAHLRRKSGSGFRGASIRYGVTRHGHQRMQARIGRAYKNDLYLTFSTOEAAE 439  
 Db 66 ATEPGRRRRK-----NVYRG-I-RKRPQKMAAEIRDRKGVRLVGFNTAEERAM 116  
 Cy 440 AYDVAALIKRGAAATNPISRYDYERIMASNLASGLARKKDNPRNDIYKSVY 499  
 Db 117 AYDVAALIKRGKAKLNPFDLHNP-----PPNYTPPSSSRSTDPQAKK-----V 163  
 Cy 500 TSYNNEETVQVQAGNNNNNENDESEWKLFNHPSQOQANNGSDQKLNCG-----NYRN 554  
 Db 164 CVV-----SQSSE-----LSQSPFVEICIGONGDEFQNLVGFEPYDL 204  
 Cy 555 SAFSMALODLLIDSVSGQHMLDES 581  
 Db 205 KQOISLSLEFLDLGNATAEQPSQLDES 231  
 RESULT 5  
 RAVI\_ARATH STANDARD; PRT; 344 AA.  
 AC O92M99;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE DNA-binding protein RAV1.  
 GN RAV1 OR ATG13360 OR Te04.2.  
 OS Arabidopsis thaliana (mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seedling;  
 RL MEDLINE=99081843; PubMed=9862967;

RA Kagaya Y., Ohmura K., Hattori T.;  
 RT "RAV1, a novel DNA-binding protein, binds to bipartite recognition  
 RT sequence through two distinct DNA-binding domains uniquely found in  
 RT higher plants";  
 RL Nucleic Acids Res. 27:470-478(1999).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RC MEDLINE=21016719; PubMed=1130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federici N.A., Kaul S.,  
 RA White O., Alonso J., Alami H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chan O., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dwar K.,  
 RA Dunn P., Egu P., Feldblum I.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Klaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kuritz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Lures J.S., Maiti R., Marziani A.,  
 RA Miltcher J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Talon L.J., Tambunga G., Tortum M.J., Town C.D.,  
 RA Utecherback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter U.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 408:816-820(2000).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RC MEDLINE=22954850; PubMed=14593172;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 RA Karlin-Newman G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Kirland-Newman G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Tortum M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Egu P., Goldsmith A.D., Gujral M., Hansen N.F.,  
 RA Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karne M.,  
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kamuya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai Y.,  
 RA Satou M., Tamez R., Vaysberg M., Wallender E.K., Wong C., Yamamura T.,  
 RA Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 RT genome";  
 RL Science 302:842-846(2003)  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RC MEDLINE=20181733; PubMed=10715325;  
 RA Fujimoto S.Y., Ohta M., Usui A., Shinishi H., Ohme-Takagi M.;  
 RT "Arabidopsis ethylene responsive element binding factors act as  
 RT transcriptional activators or repressors of GCC box mediated gene  
 RT expression";  
 RL Plant Cell 12:393-404(2000).  
 [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RC MEDLINE=21016720; PubMed=1130713;  
 RA Salanoubat M., Lemke K., Rieger M., Ansoerge M., Unseid M.,  
 RA Fattmann B., Valle G., Bioecker H., Perez-Alonso M., Obermaier B.,  
 RA Deleney M., Boutry M., Grivell L.A., Maché R., Pulidomenech P.,  
 RA De Simone V., Chotiane N., Artiguenave F., Robert C., Broctier P.,  
 RA Wincker P., Catolico L., Weissbach J., Saurin W., Querier F.,  
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,  
 RA Wumbach E., Drzonek H., Erle H., Jordan N., Bangert S.,  
 RA Wiedemann R., Kranz H., Voos H., Holland K., Brandt F., Nyakatura G.,  
 RA Vezzi A., D'Angelo W., Pallavicini A., Toppo S., Simionati B.,

DR PIR: T51329; T51329.  
 DR HSSP: 080337; 2GCG.  
 DR InterPro: IPR003340; TF\_B3.  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain1.1.  
 DR Pfam: PF02362; B3; 1.  
 DR PRINTS: PR00367; ETHRESPLENT.  
 DR PRODOM: PD001423; TF\_ERF; 1.  
 KM DNA-binding; Nuclear protein.  
 FT DNA BIND 59 121  
 FT DOMAIN 145 338  
 FT B3.  
 SQ SEQUENCE 344 AA; 38597 MW; 7349B640B3505823 CRC64.  
 Query Match 4.0%; Score 141.5; DB 1; Length 344;  
 Best Local Similarity 23.4%; Pred. No. 0.013;  
 Matches 61; Conservative 37; Mismatches 86; Indels 77; Gaps 10;  
 QY 379 SROEYVHLNRKSSGFS-----RGASIVGTRHGHGRGARGVAGN 423  
 DB 25 AKSSVGNLVRMGSSGVLDSENGVEASRLPSKYGAV-PPNGRMGMOI--YEKH 81  
 QY 424 KDYLIGFSTQEEAAYDYAAIKERGANAVYFDISRVDVERIMASSNLAGEIARRK 483  
 DB 82 QRVWLGTENEEDBAARAAYDAVAFRRRDVYVNFK---DV-----KM 120  
 QY 484 DNDPRNDIDYNSVTVSVNNEETVOVQAGNNNNENDESMKVLFNHPQOQANGSGD 543  
 DB 121 DED---EVDP---LNSHKSSEIVDMKRTNBELEGSK-----RRRNGNM 162  
 QY 544 QKIMCNGNYSASFMAUDLDIGDSVSGQHMDSESKITHTSNYSLSYSSSRE 603  
 DB 163 TRLLTSGLSND-----GVSTGFRSAELFEKAVTPSDVCKLRVVI----- 205  
 QY 604 ASPEKRGPSLLFPMPMETKI 624  
 DB 206 ---PKHHAHGFFLPSSNSV 223  
 RESULT 6  
 ERF4\_ARATH STANDARD; PRT; 222 AA.  
 AC 080340;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ethylene responsive element binding factor 4 (ATERF4).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 CX NCBI\_TaxID=3702;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20181733; PubMed=10715325;  
 RA Fujimoto S.Y., Ohta M., Usui A., Shinishi H., Ohme-Takagi M.;  
 RT "Arabidopsis ethylene responsive element binding factors act as  
 RT transcriptional activators or repressors of GCC box mediated gene  
 RT expression";  
 RL Plant Cell 12:393-404(2000).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RC MEDLINE=21016720; PubMed=1130713;  
 RA Salanoubat M., Lemke K., Rieger M., Ansoerge M., Unseid M.,  
 RA Fattmann B., Valle G., Bioecker H., Perez-Alonso M., Obermaier B.,  
 RA Deleney M., Boutry M., Grivell L.A., Maché R., Pulidomenech P.,  
 RA De Simone V., Chotiane N., Artiguenave F., Robert C., Broctier P.,  
 RA Wincker P., Catolico L., Weissbach J., Saurin W., Querier F.,  
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,  
 RA Wumbach E., Drzonek H., Erle H., Jordan N., Bangert S.,  
 RA Wiedemann R., Kranz H., Voos H., Holland K., Brandt F., Nyakatura G.,  
 RA Vezzi A., D'Angelo W., Pallavicini A., Toppo S., Simionati B.,

RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,  
 RA Reichelt J., Scharte M., Schoen O., Barges M., Terol J., Clement J.,  
 RA Navarro P., Collado C., Perez-Perez A., Oterwaelder B., Duchemin D.,  
 RA Cooke R., Landie M., Berger-Llauró C., Purnelle B., Masny D.,  
 RA de Haan M., Maarse A.C., Alcaraz J.-P., Cortic A., Casacuberta E.,  
 RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,  
 RA Mannhaupt G., Haase D., Schoof H., Rüd S., Zaccaria P., Mewes H.-W.,  
 RA Meyer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
 RA Rooney T., Rizzo C., Wals R., Utterback T., Fujii C.Y., Shea T.P.,  
 RA Creasy T.H., Haas B., Malt R., Wu D., Peterson J., Van Aken S.,  
 RA Pat G., Miltsecher J., Sellers P., Gill J.E., Feldblum T.V.,  
 RA Pruss D., Lin X., Nieman W.C., Salzberg S.L., White O., Venter J.C.,  
 RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,  
 RA Saameto S., Kimura T., Ideawa K., Kawashima K., Kishida Y.,  
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakayama S., Nakazaki N., Shimpo S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.,  
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
 thaliana.",  
 RT thaliana.",  
 RL Nature 408:820-822(2000).  
 CC -1- FUNCTION: Transcription factor that binds to the GCC-box  
 pathogenesis-related promoter element.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: Contains 1 AP2/ERF domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB008106; BAA32421.1; -  
 DR EMBL: AP000413; BAB02150.1; -  
 DR PIR: T52019; T52019.  
 DR HSSP: O80337; 2GCC.  
 DR TRANSFAC: T04636; -  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 1.  
 DR PRINTS: PR00367; ETRSPPELEMT.  
 DR PRODOM: PD001423; TF\_ERF; 1.  
 DR SMART: SM00380; AP2; 1.  
 KM Transcription regulation, DNA-binding, Nuclear protein.  
 FT DNA BIND 22 86 AP2/ERF  
 FT DOMAIN 132 137 POLY-GLY  
 SQ SEQUENCE 222 AA; 23741 MW; CEF16ABA5ACFB073 CRC64;  
 Query Match 4.0%; Score 139.5; DB 1; Length 222;  
 Best Local Similarity 27.2%; Pred. No. 0.0033;  
 Matches 67; Conservative 22; Mismatches 82; Indels 75; Gaps 11;  
 QY 401 YRGVTHHGGKQWQIGVAGKGD-LVGTSTQGEAEADVAIFRGANVTN-- 457  
 Db 25 YRGV-KRPKRGKRAAEI-KDPGKTRVWGTDTLBEAPADYADARFGKAKTNPPT 82  
 QY 458 -DISRYVERIMASSNLLAGELARRKNDPRNKDIDYNSVTVNNEETVOYAGNN 515  
 Db 83 FLELSQKVPYTFGARSP-----SQGSTLDCAPELTVVPSATA 120  
 QY 516 NNENDSFWKVVLLNHSQOQANGNSDQKIMNCGVNSRASFMLQ-----DLIGDS 569  
 Db 121 GNV-----PQLELSIGSGGG-----GSCYQIMSPVYFLDMIGN 158  
 QY 570 VSGGQHNMIDESSK-----IGTHFSNTSLVTSLSRRASPEKGPSSILPFW 617  
 Db 159 VGRGQPPVTSARFSPVAVATMAQASDSDSSVVDV---SGMEKRSCLLDLDLN 214  
 QY 618 --PME 621  
 Db 215 LPPPS 220

RESULT 7  
 ID ERF2\_ARATH STANDARD: PRT; 243 AA.  
 AC 080338;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Ethylene responsive element binding factor 2 (ATERF2).  
 GN ERF2 OR ERF-2 OR AT5G47220 OR MOLS.7.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20181733; PubMed=10715325;  
 RA Fujimoto S.Y., Ohta M., Usui A., Shinsui H., Ohme-Takagi M.;  
 RT "Arabidopsis ethylene responsive element binding factors act as  
 RT transcriptional activators or repressors of GCC box mediated gene  
 RT expression.",  
 RL Plant Cell 12:393-404(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20181125; PubMed=10718197;  
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
 RA Tsubata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
 RT feature of the regions of 3,076,755 bp covered by sixty P1 and TAC  
 RT clones.",  
 RL DNA Res. 7:31-63(2000).  
 CC -1- FUNCTION: Transcription factor that binds to the GCC-box  
 CC pathogenesis-related promoter element.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: Contains 1 AP2/ERF domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB008104; BAA32419.1; -  
 DR EMBL: AB018117; BAA97155.1; -  
 DR PIR: T51989; T51989.  
 DR HSSP: O80337; 2GCC.  
 DR TRANSFAC: T04634; -  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 1.  
 DR PRINTS: PR00367; ETRSPPELEMT.  
 DR PRODOM: PD001423; TF\_ERF; 1.  
 DR SMART: SM00380; AP2; 1.  
 KM Transcription regulation, DNA-binding, Nuclear protein.  
 FT DNA BIND 23 26 POLY-GLY  
 FT DOMAIN 114 179 AP2/ERF  
 FT DNA BIND 192 206 POLY-SER.  
 SQ SEQUENCE 243 AA; 26797 MW; C9A4C9791249B5D4 CRC64;  
 Query Match 3.8%; Score 134; DB 1; Length 243;  
 Best Local Similarity 28.6%; Pred. No. 0.026;  
 Matches 60; Conservative 30; Mismatches 82; Indels 38; Gaps 8;  
 QY 302 VTRRWKTRRYEHLWDN-----SCKKEQTRKGRQVYIGYDMEKAR--AYDLA--A 351  
 Db 17 ITRHLGGGGENELRLNSTPSSCTPS-----WGGPLKENDSEMLVGLLKDA 67  
 QY 352 LKYGSPSTHINSIENVOYLEMKMNSROEYVAHLARKSGFERGASTYRGYTRHHQHG 411  
 Db 68 FHFDTSSSDJSCLPFPRAVKVETENFTAMEEKRAKAIPTVETAVKAKHYGV-RORPWG 126



CC 412 RMOARIGRVAGN-KOLYIGTSTOEAEAYVAIKRGANVATNF-----DISR 461  
 CC 127 KFAAIEIRPAPNGARVMTGTFTEDADALADIAAFMRGSRLLNPLFVNSGEPVPR 186  
 DB 462 YDVERIMASSNLLA-----GELARRKX 483  
 DB 187 ITSXKSSSSSSSSSSSTSSSENGKXKRRK 216

RESULT 8  
 CCTL\_MOUSE STANDARD; PRT: 724 AA.  
 ID CCTL\_MOUSE  
 AC 09OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Cyclin T1 (Cyclin T) (CyclT).  
 GN CCNT1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RX MEDLINE=95059742; PubMed=9843510;  
 RA Biersatz P.D., Girdina T.A., Bogerd H.P., Cullen B.R.;  
 RT "Recruitment of a protein complex containing Tat and cyclin T1 to TAR  
 RT governs the species specificity of HIV-1 Tat.";  
 RL EMBO J. 17:7056-7065(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=95051315; PubMed=9832504;  
 RA Garber M.E., Wei P., KewalRamani V.N., Mayall T.P., Herrmann C.H.,  
 RA Rice A.P., Littman D.R., Jones K.A.;  
 RT "The interaction between HIV-1 Tat and human cyclin T1 requires zinc  
 RT and a critical cysteine residue that is not conserved in the murine  
 RT CyclT protein.";  
 RL Genes Dev. 12:3512-3527(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND MUTAGENESIS.  
 RC TISSUE=Spleen;  
 RX MEDLINE=99263519; PubMed=10329126;  
 RA Kwak Y.T., Ivanov D., Guo J., Nee B., Gaynor R.B.;  
 RT "Role of the human and murine cyclin T proteins in regulating HIV-1  
 RT Tat-activation.";  
 RL J. Mol. Biol. 288:57-69(1999).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND MUTAGENESIS.  
 RC TISSUE=Fibroblast;  
 RX MEDLINE=99145542; PubMed=9990016;  
 RA Fujinaga K., Tabbe R., Wimmer J., Cujec T.P., Peterlin B.M.;  
 RT "Interactions between human cyclin T, Tat, and the transactivation  
 RT response element (TAR) are disrupted by a cysteine to tyrosine  
 RT substitution found in mouse cyclin T.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:1285-1290(1999).  
 CC -1- FUNCTION: Regulatory subunit of the cyclin-dependent kinase pair  
 CC (CDK2/cyclin T) complex, also called positive transcription  
 CC elongation factor B (p-TEFb), which is proposed to facilitate the  
 CC transition from abortive to productive transcription by  
 CC phosphorylating the CTD (carboxy-terminal domain) of the large  
 CC subunit of RNA polymerase II (RNAP II). Does not bind efficiently  
 CC to the transactivation domain of the HIV-1 nuclear transcriptional  
 CC activator, Tat.  
 CC -1- SUBUNIT: Associates with CDK2 to form p-TEFb. Cyclin T1 is the  
 CC predominant cyclin associated with CDK2 (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. STRONG TO CYCLIN C  
 CC SUBFAMILY.

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial/  
 CC entities requires a license agreement (see <http://www.ebi-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ebi.ac.uk](mailto:license@sdb.ebi.ac.uk)).

DR EMBL: AF095640; AAD13656.1;  
 DR EMBL: AF109179; AAD19654.1;  
 DR EMBL: AF087662; AAD17798.1;  
 DR EMBL: AF113951; AAD17205.1;  
 DR MGJ: MGJ1328363; Ccmt1.  
 DR InterPro: IPR006670; Cyclin.  
 DR InterPro: IPR006671; Cyclin\_N.  
 DR Pfam: PF00134; Cyclin\_1.  
 DR SMART: SM00385; CYCLIN\_1.  
 DR PROSITE: PS00292; CYCLINS; FALSE\_NEG.  
 KW Cyclin; Cell cycle; Cell division; Coiled coil;  
 KW Transcription regulation; Nuclear protein.  
 FT DOMAIN 253 270 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 254 272 TAT-TAR RECOGNITION MOTIF (TRM).  
 FT DOMAIN 319 322 POLY-SER.  
 FT DOMAIN 384 425 COILED COIL (POTENTIAL).  
 FT DOMAIN 516 525 POLY-HIS.  
 FT DOMAIN 565 569 POLY-SER.  
 FT DOMAIN 715 723 POLY-PRO.  
 FT CONFLICT 48 48 V -> M (IN REF. 3).  
 SQ SEQUENCE 724 AA; 80565 MW; 7396E428F5A8B91B CRC64;

Query Match 3.8%; Score 134; DB 1; Length 724;  
 Best Local Similarity 19.4%; Pred No. 0.12; Indels 276; Gaps 30;  
 Matches 135; Conservative 86; Mismatches 197;

19 LGFSLS---PH-----MKKETSAAVTPPTFYMSQSGLHSFGMCYGVNGNHSPL 69  
 DB 144 LGFELTIDHHTVAVKCOLVRASKDLAQTYSFMATNSLHTTFSLOY-----TTP 194  
 QY 70 TWNPPLMSDGLCTLEALKSQNQVWPTSSPKLDELFGATNG-----THEYSGHERG 122  
 DB 195 VV-----ACVCHLCKSNWEIPIVSTGDKHWEYV-DATVLELDELTHFF----- 241  
 QY 123 LSLDSIYVNSQNAEOPNDLSQPPROQGNHVSQTHPYSGIACHGLYQAPLEETTYE 182  
 DB 242 -----LQILEKPSRLKIRINRWAY-QAAMKTPDDRQ-----ADENTTSQ 281  
 QY 183 THVSDCSLAPQMTGKXNVAPTRFSTHQVLEQMGNGMNEENGVSLSVGGEIQ 242  
 DB 282 T-----ILNMISQTSDDT-----IA 297  
 QY 243 SLISMSPGSQSCVTPASGTD-----VAVDAKRGHAKLGORPYHRSIDTFGORTSQ 298  
 DB 298 GL-MSMSTASTANVPSLPSESESSSLTSVD-MLQGBRMLSQPPFLLEAAQ--GHRTSE 353  
 QY 299 ---YRGVTRHRTGRVFAHLMDNSCKKEGGTRGRQ-----VYIGYDMERKA 344  
 DB 354 SLALIGV-----DHSLOQDGSAPGSOQASVSPASVSLKEY---RAK 395  
 QY 345 RAYVLALKTWGPETHINFSIENVOLQEMKMSRQEVYAHHRKSSGFSRGAISYRGV 404  
 DB 396 HAEELAAOKR-----OLEMNEANVSQY-----AYAAQNL 425  
 QY 405 TRHHQGRWQARIGRVAGNKDLYIGTSTOEAEAYDAVAIAFR-----GANAIVNFDI 459  
 DB 426 LSHDSHSSVILKPIESSE-----NPRFPLDXADKALKRRLPVASGDRAVSS-- 474  
 QY 460 SRVVERIMASSNLLAELARRKKNDDPNKNDIYNVSVTSVNNETVVOAGNNNN 519  
 DB 475 -----KPEIKRIRIVHAGDKHNSIDSVYKSRHKKKQKTHSNHHHH 520  
 QY 520 DSEWKMVLEHNPSSQOQ-----ANGNSDQIKINCGNYSAFSMLADLIGIDSVG 571  
 DB 521 -----NHSHRSHSLQLPAGPVSKRPEDPK----- 545  
 QY 572 SGQNMMLDESSKIGTHSNISLVTLSLSSNEAPPEKRGF----- 611

Db 546 -----HSSGOSTLAKTIVSLSSLESS--SSRRCAPRENGCAVDPHAKIAKSTR 595  
Oy 612 -SLLFPMPEMKIVNPICSTVTSULP--STVQMR 644  
Db 596 SSLLNFPPLPMTQLPQHSSDTSGLPFSQPSCKTR 631

RESULT 9  
ID CCT1\_HORSE  
ID CCT1\_HORSE STANDARD: PRT: 727 AA.  
AC Q9XTZ6; 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Cyclin T1 (Cyclin T) (CCT1).  
GN CCT1.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
NC Mammalia; Euteleota; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fibroblast;  
RX MEDLINE=99303751; PubMed=10373508;  
RA Bishiasz P.D., Girdina T.A., Bogerd H.P., Cullen B.R.;  
RT "Highly divergent lentiviral Tat proteins activate viral gene  
expression by a common mechanism";  
MoJ. Cell. Biol. 19:4592-4599(1999).  
-1- FUNCTION: Regulatory subunit of the cyclin-dependent kinase pair  
(CDK9/cyclin T) complex. Also called positive transcription  
elongation factor B (P-TEFb), which is proposed to facilitate the  
transition from abortive to production elongation by  
phosphorylating the CTD (carboxy-terminal domain) of the large  
subunit of RNA polymerase II (RNAP II) (By similarity). Binds also  
to the transactivation domain of the equine infectious anemia  
virus (EIAV) nuclear transcriptional activator, Tat. It is  
probably the cis-acting transactivation response element (TAR)  
RNA-binding cofactor for Tat. Does not bind to the transactivation  
domain of the HIV-1 nuclear transcriptional activator.  
-1- SUBUNIT: Associates with CDK9 to form P-TEFb. Cyclin T1 is the  
predominant cyclin associated with CDK9 (By similarity).  
-1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
-1- SIMILARITY: Belongs to the cyclin family. Cyclin C subfamily.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: AB137509; AAC38518.1; -  
DR InterPro: IPR006670; Cyclin.  
DR InterPro: IPR006671; Cyclin\_N.  
DR Pfam: PF00134; Cyclin\_1.  
DR SMART: SM00385; CYCLIN; 1.  
DR PROSITE: PS00292; CYCLINS; FALSE NEG.  
KW Cyclin; Cell cycle; Cell division; coiled coil;  
KW Transcription regulation; Nuclear protein.  
FT DOMAIN 253 270 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
FT 254 272 TAR-TAR RECOGNITION MOTIF (TRM) (BY  
FT SIMILARITY).  
FT FT 386 427 COILED COIL (POTENTIAL).  
FT DOMAIN 519 528 POLY-HIS.  
FT 562 573 POLY-SER.  
FT DOMAIN 718 726 POLY-PRO.  
SQ SEQUENCE 727 AA; 81013 MW; BFC2A98DB35BCE CRC64;  
Query March 3.7%; Score 129.5; DB 1; Length 727;  
Best Local Similarity 20.1%; Pred. No. 0.26; 221;  
Matches 137; Conservative 80; Mismatches 243; Gaps 31.



RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=2018125; PubMed=10718197;  
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asanizu E., Kotani H.,  
 RA Tabata S.,  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC  
 RT clones."  
 RL DNA Res. 7:31-63(2000).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=22954850; PubMed=14593172;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Maranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 RA Arakawa T., Banh U., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,  
 RA Hayashizaki Y., Johnson-Ropes C., Hsuang Y.W., Iida K., Karnes M.,  
 RA Khan S., Koeseema E., Ishida J., Jeng P.X., Jones T., Kawai J.,  
 RA Kamita A., Meyers C., Nishida J., Nishida M., Narusaka M., Seki M., Sakurai T.,  
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 RT genome."  
 RL Science 302:842-846(2003).  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,  
 RA Nakajima M., Enju A., Kamita A., Narusaka M., Carninci P., Kawai J.,  
 RA Hayashizaki Y., Shinozaki K.,  
 RT "Arabidopsis thaliana full-length cDNA."  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Transcription factor that binds to the GCC-box  
 CC pathogenesis-related promoter element.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: Contains 1 AP2/ERF domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB008107; BAA32422.1; -;  
 DR EMBL; AB018117; BAA97157.1; -; ALT\_INIT.  
 DR EMBL; AF385709; AAK60301.1; -;  
 DR EMBL; AY078804; AAL77215.1; -;  
 DR EMBL; AK117368; BAC42229.1; -;  
 DR F1R; T52020; T52020.  
 DR HSP; O80337; 2GCG.  
 DR TRANSFAC; T04637; -;  
 DR INTERPRO; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 1.  
 DR PRINTS; PR00367; ETRSPDELENT.  
 DR PRODOM; PD001423; TF\_ERF; 1.  
 DR SMART; SMD00380; AP2; 1.  
 KM Transcription regulation; DNA-binding; Nuclear protein.  
 FT DOMAIN 51  
 FT DNA\_BIND 153 218 AP2/ERF.  
 FT POLY-SER.  
 SQ SEQUENCE 300 AA; 33810 MW; 1189D464A28F7251 CRC64;  
 Query March 3.7%; Score 128.5; DB 1; Length 300;  
 Best Local Similarity 26.4%; Pred. No. 0.084;  
 Matches 55; Conservative 24; Mismatches 66; Indels 49; Gaps 11;

DB 126 LPKTEM-----IQFAEN--TKPEVTKPVSEEE-----KCHYGV-RQR 162  
 QY 409 QHGRQARI--GRVAGNDLYIGTFSQGEAAEYDVAAIKPRGANAVTNPDISRYVER 466  
 DB 163 PWGKAFAEIRDPNKKGR-VWLGTFDTRIAAARAYDEAARFKGSKALINFPLE----- 215  
 QY 467 INASNNLAGEIARRKQND--PRNKIDYKSVTSV--NNEETVQVAG-----NNNE 518  
 DB 216 -----VGKKVPRADGEGKKRRDDKVTVEKVLKTEQSVYVNGGTFPPVSNLT 267  
 QY 519 NDEWKVLE-F-NHP 531  
 DB 268 ELCDMDLGFNFP 281  
 RESULT 11  
 ID RAD9 YEAST STANDARD; PRT; 1309 AA.  
 AC P14737; O04920;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE DNA repair protein RAD9.  
 GN RAD9 OR YDR217C OR YD9534.02C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_Taxid=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89313732; PubMed=2664461;  
 RA Schiestl R.H., Reynolds P., Prakash S., Prakash L.,  
 RT "Cloning and sequence analysis of the Saccharomyces cerevisiae RAD9  
 RT gene and further evidence that its product is required for cell cycle  
 RT arrest induced by DNA damage."  
 RL Mol. Cell. Biol. 9:1882-1896(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RX MEDLINE=91061763; PubMed=2247073;  
 RA Weinert T.A., Hartwell L.H.,  
 RT "Characterization of RAD9 of Saccharomyces cerevisiae and evidence  
 RT that its function acts posttranscriptionally in cell cycle arrest after  
 RT DNA damage."  
 RL Mol. Cell. Biol. 10:6554-6564(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Murphy L., Harris D., Barrell B.G., Rajandream M.A.,  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP PHOSPHORYLATION, AND INTERACTION WITH RAD53.  
 RX MEDLINE=98429491; PubMed=9755168;  
 RA Valard J.E., Gilbert C.S., Green C.M., Lowndes N.F.,  
 RT "The budding yeast Rad9 checkpoint protein is subjected to  
 RT Mec1/Trf1-dependent hyperphosphorylation and interacts with Rad53  
 RT after DNA damage."  
 RL EMBO J. 17:5679-5688(1998).  
 CC -1- FUNCTION: Essential for cell cycle arrest at the G2 stage  
 CC following DNA damage by X-irradiation or inactivation of DNA  
 CC ligase.  
 CC -1- SUBUNIT: Physically associates with RAD53.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: Contains 1 BRCT domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M26049; AAA34954.1; -;  
 DR EMBL; Z48612; CAA88497.1; -;

[illegible]

Db	988	FSGSEIRT	-----	GNVDFKDCIFLTSLFENRELR	1017
Qy	607	E-----	KRGSLFLPMP-PMEKINIPICTS-----	YTSMP	637
Db	1018	QTIESGGTVIBSGFSLFENFTPHLAKSLVKNKNTNINRELAKLKLWMP	1066		
RESULT 12					
AC	ATRX_PANTR	STANDARD:	PRT: 2492 AA.		
AC	OTYOM4:				
DT	15-MAR-2004	(Rel. 43, Created)			
DT	15-MAR-2004	(Rel. 43, Last sequence update)			
DT	15-MAR-2004	(Rel. 43, Last annotation update)			
DE	Transcriptional regulator ATRX (X-linked helicase II) (X-linked nuclear protein) (XNP).				
GN	ATRX.				
OS	Pan troglodytes (Chimpanzee).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini, Hominiidae; Pan.				
RN	NCBI_TaxId=9598;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=22763540; PubMed=12777533;				
RA	Kitano T., Schwarz C., Nickel B., Pabo S.;				
RT	"Gene diversity patterns at 10 X-chromosomal loci in humans and chimpanzees."				
RL	Mol. Biol. Evol. 20:1281-1289(2003)				
CC	-1- FUNCTION: Could be a global transcriptional regulator. Modifies gene expression by affecting chromatin. May be involved in brain development and facial morphogenesis.				
CC	-1- SUBUNIT: Probably binds E2H2. Binds annexin V in a calcium and phosphatidylcholine/phosphatidylserine-dependent manner (By similarity).				
CC	-1- SUBCELLULAR LOCATION: Nuclear. Associated with pericentromeric heterochromatin during interphase and mitosis, probably by interacting with HPI.				
CC	-1- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.				
CC	-1- SIMILARITY: Contains 1 PHD-type zinc finger.				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-slb.ch/announce/">http://www.isb-slb.ch/announce/</a> or send an email to <a href="mailto:license@isb-slb.ch">license@isb-slb.ch</a> ).				
CC	-----				
DR	EMBL; AB102642; BAC8111.1; -				
DR	InterPro; IPR001410; DEAD.				
DR	InterPro; IPR001650; Helicase_C.				
DR	InterPro; IPR000330; SNF2_N.				
DR	Pfam; PF003271; helicase_C_1.				
DR	Pfam; PF00176; SNF2_N_1.				
DR	SMART; SM00487; DEXDC; 1.				
DR	SMART; SM00490; HELICC; 1.				
KW	DNA repair; Nuclear protein; DNA-binding; Helicase; ATR-binding; Zinc-finger.				
FT	NP_FING	220	268		
FT	NP_BIND	1594	1691		
FT	SITE	1719	1722		
FT	DOMAIN	745	750		
FT	DOMAIN	1151	1156		
FT	DOMAIN	1166	1169		
FT	DOMAIN	1202	1206		
FT	DOMAIN	1259	1266		
FT	DOMAIN	1443	1466		
FT	DOMAIN	1499	1502		
FT	DOMAIN	1929	1939		
FT	DOMAIN	1941	1948		
FT	DOMAIN	2232	2235		
FT	DOMAIN	2262	2265		
FT	DOMAIN				
FT	PHD-TYPE.				
FT	ATP (POTENTIAL).				
FT	DEGR_BOX.				
FT	POLY-SRR.				
FT	POLY-SRR.				
FT	POLY-SRR.				
FT	POLY-ASP.				
FT	POLY-GLU.				
FT	POLY-GLU.				
FT	POLY-LYS.				
FT	POLY-SRR.				
FT	POLY-LYS.				
FT	POLY-GLU.				



```

y 544 OKIMNCGANTYNSAFSMALQDILIGDVSQGGQHNMLDESSXIGTGHFSTSLTSSSRE 603
Db 518 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 604 ASPEK--RGPBLPEPKMPMETKTVNPIGTSVTSMLPSPVQMRSPAISLSHLVPFASKT 661
Db 578 RSPHSHINGPMLT---PSSSTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
RESULT 14
CCT1_PANTR STANDARD. PRT: 725 AA.
AC 08HXN7;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cyclin T1 (Cyclin T) (CyclT).
OS CNNT1.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_Taxid=9598;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=23328476; PubMed=12441807;
RX Rollman E.N., Lund L.H., Sjoststrand D.E., Leitner T., Wahren B.E.;
FA "A unique amino acid deletion in the chimpanzee cyclin T1 does not
RT affect Tat trans-activation of HIV."
RL AIDS 16:2335-2337(2002).
CC -1- FUNCTION: Regulatory subunit of the cyclin-dependent kinase pair
CC (CDK9/cyclin T) complex, also called positive transcription
CC elongation factor B (P-TEFb), which is proposed to facilitate the
CC transition from abortive to production elongation by
CC phosphorylating the CTD (carboxy-terminal domain) of the large
CC subunit of RNA polymerase II (RNAP II). Binds also to the
CC transactivation domain of the HIV-1 nuclear transcriptional
CC activator, Tat. It is probably the cis-acting transactivation
CC response element (TAR) RNA-binding cofactor for Tat. Also serves
CC as an essential cofactor for HIV 2-Tat and simian immunodeficiency
CC virus (African mandrill) Tat function.
CC -1- SUBUNIT: Associates with CDK9 to form P-TEFb. Cyclin T1 is the
CC predominant cyclin associated with CDK9 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the cyclin family. Cyclin C subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AF520776; AAN73282.1; -
CC InterPro: IPR006670; Cyclin.
CC DR InterPro: IPR006671; Cyclin_N.
CC DR Pfam: PF00134; cyclin_1.
CC DR SMART; SM00385; CYCLIN; 1.
CC DR PROSITE; PS00292; CYCLINS; FALSE NEG.
CC DR Cyclin; Cell cycle; Cell division; Colled coil;
CC Transcription regulation; Nuclear protein.
CC KW DOMAIN 252 269
CC FT DOMAIN 253 271 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT DOMAIN 383 424 TAR-TAR RECOGNITION MOTIF (TRM).
CC FT DOMAIN 316 329 COILED COIL (POTENTIAL).
CC FT DOMAIN 539 529 POLY-HIS.
CC FT DOMAIN 559 569 POLY-SER.
CC FT DOMAIN 716 724 POLY-PRO.
CC SQ SEQUENCE 725 AA; 80568 MW; A21C398F6A4378A1 CRC64;

Query Match 3.5%; Score 121; DB 1; Length 725;
Best Local Similarity 20.1%; Pred. No. 0.99;
Matches 137; Conservative 76; Mismatches 220; Indels 250; Gaps 31

```

QY	1	LGSGS---	PH-----	MMKATSAATVPTTFVYSPQSHLSNFGMVGVEGNGNHSPL	66
Db	143	LGFEELLIDHPHHVVKCTQLVRASKDLAQISYFMAATNSLHLTFFSLQY	-----	TPP	193
QY	70	TYWPELKSQSGSLCILEALRKSGTQVMVPPSPSKLEDIFLGATMG	-----	THEYSGHERG	122
Db	194	VV-----	ACVCHILHACKWMSWMEIPVSDGKHWEYV-DATVTLLELDELTHF	-----	240
QY	123	LSLDSITYNSQAABQPNRDLISGPFROQGHMSVQTHPYVSGLACHGLYOAPLEETPK	-----	182	
Db	241	-----	LQLEETPR-----	LKINWTRAC	265
QY	183	THVSDGSSIMPQMTBGLKNWAPLPRFESTHOQVLEQQMCGMNGERNGVLSGVCGLQ	-----	242	
Db	266	TKADD-----	RGDDEKTSBOTIL-----	NMIS-----	288
QY	243	SLSLMSPGSSQCVTAPSGTDSYVADKKKGHAKLGOQOPVHR	-----	KSIDTFQG	294
Db	289	SSSDITITAGLMNMTSTTSNAPSLPVSSESSNLTSYMLPGKRWLSQSPFLPEPTQGH	-----	348	
QY	295	RISQYGVTRHRTGTRYEAHLMDNSCKTEG	-----	QTRKROYLGGYDME	347
Db	349	RTSENIALTG-----	YDHLSPDGGNAPISQKNSKVSFAKVLKEVRAKHA	357	
QY	348	DLAALKXMPSTHINPSTENYVOLEEMKMSRQEVYAHLRKSSGSRGASIVRGVTRH	-----	407	
Db	398	ELAAQR-----	QLENNEANVQY-----	AYAAQNLSH	427
QY	408	HQGRWQATIGRVAGNKLTL	-----	GTFSQOEAAEAYDVAIKFR	457
Db	428	HD-----	SHSSVILKMPTEGSENERPFLKADYTAIKRI	PVAGDPKA	471
QY	458	DISRYVERKIMASSNLNAGELLARREKNDPRNKOIDYKKSVPVTSVNNNE	-----	TVQVAGANN	517
Db	472	-----	ASSK--PEELMRLKVAADKANSVEDSTYKREHEKFKHPSNHH	518	
QY	518	ENDSEWQVTLFNHPSQCCQANGSDOKIMNCGYRNSAFSMALQDLDIGDSVSGQHNM	-----	577	
Db	519	HH-----	NHSHK--HSHSQPVGTGNR	PDGPH	546
QY	578	LDSSSKIGHFSTSLVTSLSSSSEAPBEKGRS	-----	LLFMPMETKIVAPIGTSV	632
Db	547	SSQTSNLAKHTYSLSSFSSS--SSTRKRGSEETGAVDFHPAKIAKTR	-----	SSSL	599
QY	633	TSWLPS-PTVOKRPPAISLSHL	-----	654	
Db	600	NFSPPLPIMACMPHGHSPTGSL	-----	622	
RESULT 15					
APC_HUMAN					
ID	APC_HUMAN	STANDARD;	PRT;	2843	AA.
AC	P25054; O15162; O15163;				
DT	01-MAY-1992 (Rel. 22, Created)				
DT	01-MAY-1992 (Rel. 22, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
DE	Adenomatous polyposis coli protein (APC protein).				
GN	APC OR DP2.5.				
OS	Human sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
NCBI_TaxId	9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=91335210; PubMed=1651562;				
RA	Kinzler K.W., Nilbert M.C., Su L.-K., Vogelstein B., Bryan T.M.,				
RA	Levy D.B., Smith K.J., Preisinger A.C., Hedge P., McKechnie D.,				
RA	Finnlater R., Markham A., Groffen U., Boguski M.S., Altschul S.F.,				
RA	Horii A., Ando H., Miyoshi Y., Miki Y., Nishisho I., Nakamura Y.,				
RT	"Identification of FAP locus genes from chromosome 5q21."				
RL	Science 253:661-665(1991).				
RN	[2]				

RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=fetal brain;  
 RX MEDLINE=91330307; PubMed=1678319;  
 RA Joslyn G., Carlson M., Thilveris A., Albertsen H., Gelbert L.,  
 RA Samowitz W., Gorden J., Stevens J., Spirio L., Robertson M.,  
 RA Sargent L., Krapcho K., Wolff E., Burt R., Hughes J.P.,  
 RA Warrington J., McPherson J.D., Wamuth R., Le Paslier D.,  
 RA Abderrahim H., Cohen D., Leppert W., White K.,  
 RT "Identification of deletion mutations and three new genes at the  
 RT familial polyposis locus.";  
 RL Cell 66:601-613(1991).  
 RN [3]  
 RP ASSOCIATION WITH CATENINS.  
 RX MEDLINE=94082295; PubMed=8259519;  
 RA Su L.-K., Vogelstein B., Kinzler K.W.,  
 RT "Association of the APC tumor suppressor protein with catenins.";  
 RL Science 262:1734-1737(1993).  
 RN [4]  
 RP DISEASE.  
 RX MEDLINE=95174843; PubMed=7661930;  
 RA Hamilton S.R., Liu B., Parsons R.E., Papadopoulos N., Jen J.,  
 RA Powell S.M., Krush A.J., Berk T., Cohen Z., Tetu B., Burger P.C.,  
 RA Wood P.A., Taqi F., Booker S.V., Petersen G.M., Offenshaw G.J.A.,  
 RA Termette A.C., Giardiello F.M., Vogelstein B., Kinzler K.W.,  
 RT "The molecular basis of Turcot's syndrome.";  
 RL New Engl. J. Med. 332:839-847(1995).  
 RN [5]  
 RP DISEASE.  
 RX MEDLINE=97094176; PubMed=8940264;  
 RA Eccles D.M., van der Luijt R.B., Breukel C., Bullman H., Bunyan D.,  
 RA Fisher A., Barber J., du Boulay C., Pimrose J., Burn J., Fodde R.,  
 RT "Hereditary desmoid disease due to a frameshift mutation at codon 1924  
 RT of the APC gene.";  
 RL Am. J. Hum. Genet. 59:1193-1201(1996).  
 RN [6]  
 RP DISEASE.  
 RX MEDLINE=20243021; PubMed=10782927;  
 RA Couture J., Mitrì A., Lagace R., Saults R., Berk T., Bouchard H.-L.,  
 RA Fodde R., Altman B., Bapat B.,  
 RT "A germline mutation at the extreme 3-prime end of the APC gene  
 RT results in a severe desmoid phenotype and is associated with  
 RT overexpression of beta-catenin in the desmoid tumor.";  
 RL Clin. Genet. 57:205-212(2000).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 2-55.  
 RX MEDLINE=20384842; PubMed=10926498;  
 RA Day C.L., Alber T.,  
 RT "Crystal structure of the amino-terminal coiled-coil domain of the  
 RT APC tumor suppressor.";  
 RL J. Mol. Biol. 301:147-156(2000).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 1021-1035 IN COMPLEX WITH  
 CNINB1.  
 RX MEDLINE=21564054; PubMed=11707392;  
 RA Eklof Spink K., Fridman S.G., Weis W.I.,  
 RT "Molecular mechanisms of beta-catenin recognition by adenomatous  
 RT polyposis coli revealed by the structure of an APC-beta-catenin  
 RT complex.";  
 RL EMBO J. 20:6203-6212(2001).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (1.90 ANGSTROMS) OF 2034-2049 IN COMPLEX WITH  
 AXIN.  
 RX MEDLINE=20271867; PubMed=10811618;  
 RA Spink K.E., Polakis P., Weis W.I.,  
 RT "Structural basis of the axin-adenomatous polyposis coli  
 RT interaction.";  
 RL EMBO J. 19:2270-2279(2000).  
 RN [10]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=94154728; PubMed=8111410;  
 RA Nagase H., Nakamura Y.,  
 RT "Mutations of the APC (adenomatous polyposis coli) gene.";  
 RL Hum. Mutat. 2:425-434(1993).

RN [11]  
 RP VARIANTS FAP.  
 RX MEDLINE=91335211; PubMed=1651563;  
 RA Nishio I., Nakamura Y., Miyoshi Y., Miki Y., Ando H., Horii A.,  
 RA Koyama K., Utsunomiya J., Baba S., Hedge P., Markham A., Krush A.J.,  
 RA Petersen G.M., Hamilton S.R., Nilbert M.C., Levy D.B., Bryan T.M.,  
 RA Preisinger A.C., Smith K.J., Su L.-K., Kinzler K.W., Vogelstein B.,  
 RT "Mutations of chromosome 5q21 genes in FAP and colorectal cancer  
 RT patients.";  
 RL Science 253:665-669(1991).  
 RN [12]  
 RP VARIANTS FAP.  
 RX MEDLINE=93265030; PubMed=1338904;  
 RA Miyoshi Y., Nagase H., Ando H., Ichii S., Nakatsuru S., Aoki T.,  
 RA Miki Y., Mori T., Nakamura Y.,  
 RT "Somatic mutations of the APC gene in colorectal tumors: mutation  
 RT cluster region in the APC gene.";  
 RL Hum. Mol. Genet. 1:229-233(1992).  
 RN [13]  
 RP VARIANTS FAP.  
 RX MEDLINE=93244793; PubMed=1338691;  
 RA Nakatsuru S., Yanagisawa A., Ichii S., Tanara E., Kato Y.,  
 RA Nakamura Y., Horii A.,  
 RT "Somatic mutation of the APC gene in gastric cancer: frequent  
 RT mutations in very well differentiated adenocarcinoma and signet-ring  
 RT cell carcinoma.";  
 RL Hum. Mol. Genet. 1:559-563(1992).  
 RN [14]  
 RP VARIANTS FAP TRP-1348, AND VARIANTS ASP-1118; MET-1292; VAL-1304 AND  
 SER-2502  
 RX MEDLINE=93250848; PubMed=1338764;  
 RA Nagase H., Miyoshi Y., Horii A., Aoki T., Petersen G.M.,  
 RA Vogelstein B., Maher E., Ogawa M., Maruyama M., Utsunomiya J.,  
 RA Baba S., Nakamura Y.,  
 RT "Screening for germline mutations in familial adenomatous polyposis  
 RT patients: 61 new patients and a summary of 150 unrelated patients.";  
 RL Hum. Mutat. 1:467-473(1992).  
 RN [15]  
 RP VARIANTS FAP TRP-99.  
 RP TISSUE=Peripheral blood lymphocytes;  
 RX MEDLINE=95134544; PubMed=7831149;  
 RA Dobbie Z., Spycher M., Huerliman R., Ammann R., Ammann T., Roth J.,  
 RA Mueller A., Mueller H., Scott R.J.,  
 RT "Mutation analysis of the first 14 exons of the adenomatous  
 RT polyposis coli (APC) gene.";  
 RL Eur. J. Cancer 30A:1709-1713(1994).  
 RN [16]  
 RP VARIANTS FAP GLY-722.  
 RX MEDLINE=95135430; PubMed=7831931;  
 RA Stella A., Montero M., Resta N., Marchese C., Susca F., Gentile M.,  
 RA Romio L., Pilia S., Prete F., Varenti C., Guanti G.,  
 RT "Four novel mutations of the APC (adenomatous polyposis coli) gene in  
 RT FAP patients.";  
 RL Hum. Mol. Genet. 3:1687-1688(1994).  
 RN [17]  
 RP ERRATUM.  
 RA Stella A., Montero M., Resta N., Marchese C., Susca F., Gentile M.,  
 RA Romio L., Pilia S., Prete F., Varenti C., Guanti G.,  
 RL Hum. Mol. Genet. 3:1918-1918(1994).  
 RN [18]  
 RP VARIANTS FAP ILE-171.  
 RX MEDLINE=97144176; PubMed=8990002;  
 RA van der Luijt R.B., Meera Khan P., Vaseen H.F.A., Tops C.M.J.,  
 RA van Leeuwen-Cornelisse I.S.J., Wijnen J.T., van der Klift H.M.,  
 RA Plug R.J., Griffioen G., Fodde R.,  
 RT "Molecular analysis of the APC gene in 105 Dutch kindreds with  
 RT familial adenomatous polyposis: 67 germline mutations identified by  
 RT DGGE, PTT, and southern analysis.";  
 RL Hum. Mutat. 9:7-16(1997).  
 RN [19]  
 RP VARIANTS COLORECTAL CARCINOMA THR-880; ILE-890 AND VAL-1508.  
 RX MEDLINE=98080146; PubMed=9419797;  
 RA Miyaki M., Nishio I., Konishi M., Kikuchi-Yanoshiba R., Tanaka K.,

RA Muraoka M., Nagato M., Chong J.-M., Koike M., Terada T., Kawahara Y.,  
 RA Fukutome A., Tomiyama J., Chuganji Y., Momoi M., Utsunomiya J.,  
 RA "Dietary genetic instability of tumors and normal tissues in Turoc  
 RT syndrome.";  
 RL Oncogene 15:2877-2881(1997).  
 RN [20]  
 RP VARIANT LYS-1307.  
 RX MEDLINE=98400248; PubMed=9731522;  
 RA Redaon M., Nathanson K.L., Yuan Z.Q., Neuhansen S.L., Satagopan J.,  
 RA Wong N., Yang D., Nafa D., Abrahamson J., Orzellik H.,  
 RA Antin-Ozerkis D., Andrulis I., Daly M., Pinsky L., Schrag D.,  
 RA Gallinger S., Kaback M., King M.-C., Woodage T., Brody L.C.,  
 RA Godwin A., Warner E., Weber B., Foulkes W., Offit K.;  
 RT "The APC 11307K allele and breast cancer risk.";  
 RL Nat. Genet. 20:13-14(1998).  
 RN [21]  
 RP VARIANTS LYS-1307 AND GLN-1317.  
 RC TISSUE=Peripheral blood;  
 RX MEDLINE=98393712; PubMed=9724771;  
 RA Freyling I.M., Beck N.E., Ilyas M., Dove-Edwin I., Goodman P.,  
 RA Pack K., Bell J.A., Williams C.B., Hodgson S.V., Thomas H.J.W.,  
 RA Talbot I.C., Bodmer W.F., Tomlinson I.P.M.;  
 RT "The APC variants 11307K and 11317Q are associated with colorectal  
 tumors, but not always with a family history.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:10722-10727(1998).  
 RN [22]  
 RP VARIANT LYS-1307.  
 RX MEDLINE=98400259; PubMed=9731533;  
 RA Woodage T., King S.M., Macholder S., Hartge P., Struwing J.P.,  
 RA McDanna M., Laken S.J., Tucker M.A., Brody L.C.;  
 RT "The APC 11307K allele and cancer risk in a community-based study of  
 RT Ashkenazi Jews.";  
 RL Nat. Genet. 20:62-65(1998).

Query Match 3.5%; Score 121; DB 1; Length 2843;

Best Local Similarity 19.6%; Pred. No. 7.1;  
 Matches 131; Conservative 97; Mismatches 268; Indels 174; Gaps 30;

QY 58 GVGNGNFHSPVTW--PLKSDGSLCTIEALKRSQTVMTPT--SSPKLEDFLGATMGTH 114  
 DB 2235 GVRNSSSTSPVSKKGPDKTPAS---KSPSEQITATSPRGAKPSKSELSFVARQTS 2290  
 QY 115 EYGSHERGLSDSTYYNQNAEAPNRDLISQPRCOGKMSVOTHPYSGLA--CHGLYQA 173  
 DB 2291 QIGSSSKAPSRSG--SRDSTPSRPAQPLSRPIOSPRNSIS--PGRNGISPPNKLSQL 2345  
 QY 174 PLEETTYETHVSDSSIMPMQTEGLKNVAPTEFSTHOVLEQVANC----- 222  
 DB 2346 P---RTSSPSTASTKS-----GSGKMSYTPGRQMS--QNLTKQTGLSKNMSIRSE 2395  
 QY 223 ---GNGERNGVSLGSGCELOSLISMSPGSSSCVTPASGTDVAADAKRGHAKL 278  
 DB 2396 SASKGLNOMNNG--NGAKKVELSRMSSTKSGSESRSERP-----VIVROSTPIKE 2446  
 QY 279 GOKQPVHRKSIDTFQRTSOYRGVTRHRWTRGYEAHLMDNSCKKEGQTRKGRQVYLGVD 338  
 DB 2447 APSFTLRKLL-----EESASFESLS-----PSRPASPTRSQAQTPVLSPS 2487  
 QY 339 MEKKAARAYDLAALKYMGPSHTINISINYQ-----VOLBEMKMSRQEVVAHLRR 389  
 DB 2488 LPDMGL-----STHSSVOAGWRKLPNNLSPTLEINDGRPAKSHD---IAR 2530  
 QY 390 KSGSPSRGASIRYGVTRHGRWQARIGRAV---GNKDLVLTGTFSTOEAABAYD-- 442  
 DB 2531 SHSEBSPRLPIRSGTMRKSHKSSSLPRVSTWRTGSSSSILSASSESSEKASEDEK 2590  
 QY 443 ---VAATFRGANAYTNFDIRYVERIMASSNLLAGELARKKNDPRNKDIDYNSV 498  
 DB 2591 HYNISIGTKOSKENOVSAKGTWRKI KENEFSPTN--STQTVSSGATNGAESKTLIYOMAP 2649  
 QY 499 VTS-----VNEETVOVQAGN-----NNNENDESEMRQVLFNHPDQ 533  
 DB 2650 AVSKTEDVAVRIEDCGINNPRGRSPTGTPTFVIDSVSEKANPNIKDCK-----DNQ 2701

QY 534 QOQANGNS-DOKIMNGYRNSAFSVALQDLIGDSVSGGQHNMLEDSSKIGTHFNTS 592  
 DB 2702 AKQVNGSVPMRTVGLNRLNSFIDVADPDQKQTE-IRKQNNPVVSE-----TNES 2754  
 QY 593 GLV-----TSSSRERASPEKRGPSLFPMPPMETK-VNPIGTSTV--SWLPSPTVQMRP 645  
 DB 2755 SIVERTPPSSSSSKSHSSPS-----GTVAARVTPFNINPSF-----RK 2792  
 QY 646 SPAISLSHIP 655  
 DB 2793 SSADSTSRP 2802

Search completed: March 9, 2004, 10:46:04  
 Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2004, 10:41:58 ; Search time 45 Seconds  
(without alignments)

4648.635 Million cell updates/sec

Title: US-10-024-632-2

Perfect score: 3502 1 MKRINSSNTDGNNNHNLG.....RSPPAISLHLPVPSWDT 663

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

SPTREMBL.25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacterioph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1370.5	39.1	639	10	Q84Z02
2	1336	38.1	588	10	Q885W7
3	1328	37.9	555	10	Q38914
4	1328	37.9	555	10	Q42462
5	956	27.3	692	10	Q881E1
6	953	27.2	584	10	Q8LSN1
7	947	27.0	579	10	Q8LSN2
8	942	26.9	579	10	Q8LSN3
9	923.5	26.4	425	10	Q9CAH3
10	919.5	26.3	597	10	Q8LGG3
11	918.5	26.2	555	10	Q7XPD4
12	916.5	26.2	581	10	Q9LFR2
13	900.5	25.7	555	10	Q9F1E2
14	893.5	25.5	540	10	Q9L737
15	890	25.4	485	10	Q41832
16	879	25.1	516	10	Q9SYC2

17	841.5	24.0	492	10	Q8LGG4	Q81G54	oryza sativ
18	839.5	24.0	437	10	Q9LSM4	Q91EM4	arabidopsis
19	831	23.7	566	10	Q91X43	Q91X43	arabidopsis
20	688.5	19.7	498	10	Q7XTE2	Q7XTE2	oryza sativ
21	672	19.2	345	10	Q94AN4	Q94AN4	arabidopsis
22	645	18.4	332	10	Q9S9M9	Q9S9M9	arabidopsis
23	616	17.6	308	10	Q9MA06	Q9MA06	arabidopsis
24	566	16.2	223	10	Q8RV98	Q8RV98	oryza sativ
25	534	15.2	466	10	Q7XPM5	Q7XPM5	oryza sativ
26	529	15.1	199	10	Q8GV25	Q8GV25	arabidopsis
27	517	14.8	205	10	Q9M2V4	Q9M2V4	arabidopsis
28	515.5	14.7	533	10	Q9AX14	Q9AX14	pisum sativ
29	498.5	14.2	505	10	Q84U24	Q84U24	antirrhinum
30	497	14.2	519	10	Q9XHD4	Q9XHD4	petunia hyb
31	496	14.2	422	10	Q22949	Q22949	arabidopsis
32	495.5	14.1	505	10	Q84U23	Q84U23	antirrhinum
33	495	14.1	423	10	Q8LA00	Q8LA00	arabidopsis
34	487	13.9	372	10	Q84TV8	Q84TV8	oryza sativ
35	487	13.9	434	10	Q84TB5	Q84TB5	oryza sativ
36	484	13.8	549	10	Q8VWM8	Q8VWM8	malus domes
37	483.5	13.8	428	10	Q8GWR2	Q8GWR2	arabidopsis
38	483	13.8	446	10	P93632	P93632	zea mays (m
39	477.5	13.6	436	10	Q8H443	Q8H443	oryza sativ
40	477	13.6	440	10	Q8VWQ6	Q8VWQ6	hordeum vul
41	475.5	13.6	352	10	Q9F935	Q9F935	arabidopsis
42	474.5	13.5	633	10	Q9FV33	Q9FV33	plcea abies
43	473.5	13.5	531	10	Q9FV34	Q9FV34	plcea abies
44	467.5	13.3	433	10	Q65008	Q65008	zea mays (m
45	462	13.2	457	10	Q9XHD3	Q9XHD3	petunia hyb

## ALIGNMENTS

## RESULT 1

ID Q84Z02 PRELIMINARY; PRT; 639 AA.

AC 01-JUN-2003 (TREMBLrel. 24, Created)  
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Putative ovule development protein antilegumenta (ANT).  
GN P0474G09.16.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OC NCBI\_TaxID=39947;  
OX [1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RA Sasaki T., Matsumoto T., Katayose Y.,  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC  
clone: P0474G09.16";  
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
RU EMBL:AP005309; BAC56815.1; -  
DR GO; GO:0005634; Cytosol; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001471; TF\_ERF.  
DR Pfam; PF00847; AP2-domain; 2.  
DR PRINTS; PR00367; ETRSPLEMMT.  
DR ProDom; PD001423; TF\_ERF; 2.  
DR SMART; SM00380; AP2; 2.  
DR SQUONCE 639 AA; 67622 NM; EB4BC5123C56E68F CRC64;

Query Match 39.1%; Score 1370.5; DB 10; Length 639;

Best Local Similarity 46.6%; Pred. No. 1.6e-90;

Matches 337; Conservative 68; Mismatches 169; Indels 149; Gaps 30;

QY 4 INESNTDGN---HNWIGFSLPMKWEATSAATVPTFTW-----SPGSH-- 49  
DB 1 MTNSNNGNGTAAASGWLGSLSFPM-----ASTMDHHVHHHQQQQQQQQQQHHQ 55



QY 50 -----LSNFGWCY-----VGENGNFHSPLTVMPLKSDGSLTLEALKRSQTO 92  
 DB 56 QOQHLLFPSTVTTAAAAAAGTAGVVAATNGYSQASLMSPLKSDGSLTLEALKRSQTO 115  
 QY 93 VMTPTSPKLEDFLGATMGTHYSGHSGSLD---STYNSQNAEAPKPRDLSQPR 149  
 DB 116 ---HHGPKLEDFLGAA-----QPMALSLDNTSSFYGGGGAAGH-----G 155  
 QY 150 OQGHMSVOTHPYSGTAGHGYOAPLEBETKETHVSCSSSLMPOMTEGLKNVAPTRF 209  
 DB 156 QHGYL--QACDLYGPPAPSLVTADEAAA-----AAAAASWTAAGAA 199  
 QY 210 STHQVLEQONNCGMNERNGVSLGSGCEL--QSLSLSPSGSQSCVTAAPS--GTDVA 267  
 DB 200 TAY-----ATGAADANAENVLPATAOHLHPLALMSGSGSLSCITAGEYGAAYA 253  
 QY 268 V--DAKRGHA-KIGOKOPV--HRKSIDFGORTSGYRGVTRHRTGRYEAHLMDNSCKE 323  
 DB 254 AADGKRKGAGGAGGQKOPVHHRKSIDFGORTSGYRGVTRHRTGRYEAHLMDNSCKE 313  
 QY 324 GQTRKGRQVYLGGYDMEERKAAAYDLAALKYGPSTHINFSIENVQVLEEMKMSRQY 383  
 DB 314 GQTRKGRQ---GGYDMEERKAAAYDLAALKYGPSTHINFLPDYQELIEEMKMSRQY 370  
 QY 384 VAHLRRKSGSGFRGASITRGVTRHRTGRYEAHLMDNSCKE 443  
 DB 371 VAHLRRKSGSGFRGASITRGVTRHRTGRYEAHLMDNSCKE 430  
 QY 444 AAIKRGAAVATNPDISRYDVERIMASSNLAGEIARRKNDPBNKDIDYNSKSVT 503  
 DB 431 AAIKRGAAVATNPDISRYDVERIMASSNLAGEIARRKNDPBNKDIDYNSKSVT 477  
 QY 504 NEET-----VOYQAGNNNENDESEKMTLVNHPSCQOQANG---NSDQKIMCGYR 553  
 DB 478 DHSAGAGAIIVSHSAADIAVADTLMKAT--TAPRQOQOHHDDVVLGADQ----- 526  
 QY 554 NSAFMSALODLIGDVSQGNHMLDESEKITHSNTSLVTSLSRREASPEKRG-S 612  
 DB 527 -AAFSV-LHDLVAVDAAAHAQOQOQHSM---SAASLVTLSLSNRREGSPDRGGLS 580  
 QY 613 LLEPMP-----PMETKIVN--PFGISVTWLPSPYQMPSPALISLHPVVAS 659  
 DB 581 MFAKSPAVVAASAOQAATKLAAPLPLG---SWVSSPPASARP--PAVSIAHVPFPA 635  
 QY 660 WTD 662  
 DB 636 WTD 638

RESULT 2  
 Q885W7 PRELIMINARY; PRT; 588 AA.

AC Q885W7;  
 DT 01-JUN-2002 (TEMBLrel. 21, Created)  
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)  
 DE Putative ovule development protein antilegumenta (ANT).  
 GN CUI175C11.21 OR CUI1017C11.11.  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Eukaryota; Magnoliopsida; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wang R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,  
 RT "Rice Genomic Sequence."  
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 RP [2]  
 RA Wang R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Currie J.,

RA Collura K.,  
 RT "Rice Genomic Sequence."  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DB EMBL; AC103891; AA019141.1;  
 DB EMBL; AC135157; AA00690.1;  
 DR Gramene; Q885W7;  
 DR GO; GO:0005634; Cytosol; IEA.  
 DR GO; GO:0003700; P:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 2.  
 DR PRINTS; PR00367; ETHRSPLEMT.  
 DR Prodom; PD001423; TF\_ERF; 2.  
 DR SMART; SM00380; AP2; 2.  
 DR SMART; SM00380; AP2; 2.  
 SQ SEQUENCE 588 AA; 63230 KM; A13087F5A36CEDIC CR664;  
 Query Match 38.1%; Score 1336; DB 10; Length 588;  
 Best Local Similarity 47.5%; Pred. No. 4; 3e-88;  
 Matches 333; Conservative 62; Mismatches 146; Indels 160; Gaps 30;  
 QY 7 SNTDGNHNNHGLFSLTHMKKATSAATVPTTFMSPSQSHLSNFGWCYGVGN--GN 64  
 DB 2 SSPDATTGNNFL-FS-PPAAQVAPV---PGYIV-----GAYGQGTSTAV 45  
 QY 65 FHSPLTVMPLKSDGSLTLEALKRSQTVMPVTPSSFKEDPLGATMGTHYSGHSGSL 124  
 DB 46 YSHLPVMPILKSDGSLTLEALKRSQTVMPVTPSSFKEDPLGATMGTHYSGHSGSL 87  
 QY 125 LBSITVNSQNAEAPKPRDLSQPRQGMASVQTHYVYGLACHGYOAPLEBETKETH 184  
 DB 88 -DPATYSGQOEED-----ASRAAYOHQV--PYN-----YQPLTEAMTQEA 130  
 QY 185 VSDCSLMPOMTEGLKNVAPTRFSTHQQVLEQONNCGMNERNGVSLGSGCEL--QS 243  
 DB 131 AAMPEDAM-----AAAKNF-----LVTSAGAYGNQ-----EMGP 161  
 QY 244 LSLSPSGSQ--SSCVT--APSGTDSVAVDA-----KKRGAKG 279  
 DB 162 LSLSPSGSQSSSCVSAAPQOQNAVAAAAAGDQGSNSNDGEGQVGRKRGKGG 221  
 QY 280 QKOPVHRKSIDFGORTSGYRGVTRHRTGRYEAHLMDNSCKEGRQTRKGRQVYLGGYD 339  
 DB 222 QKOPVHRKSIDFGORTSGYRGVTRHRTGRYEAHLMDNSCKEGRQTRKGRQ---GGYDT 278  
 QY 340 EERAAAYDLAALKYGPSTHINFSIENVQVLEEMKMSRQYVAHLRRKSGFRGAS 399  
 DB 279 EDKRAAYDLAALKYGPSTHINFSIENVQVLEEMKMSRQYVAHLRRKSGFRGAS 338  
 QY 400 IYRGVTRHRTGRYEAHLMDNSCKEGRQTRKGRQVYLGGYD 459  
 DB 339 IYRGVTRHRTGRYEAHLMDNSCKEGRQTRKGRQVYLGGYD 388  
 QY 460 SRYDVERIMASSNLAGEIARRKNDPBNKDIDYNSKSVTAVNNEETVOAG 513  
 DB 399 TRYDVOKIMSSSLTGEAARKYKALEAPDHPVIGREL-----GATEEASATVVG 450  
 QY 514 NNNNENDESEKMTLVNHPSCQOQANG---NSDQKIMCGYRNASAFMALODLIGDS 569  
 DB 451 -----TDMRWVL--HGSQOQOAACTEATADLQK---GWMGDAHS--ALHGIVGFV 495  
 QY 570 VGSQGNHMLDESEKITHSNTSLVTSLSRREASPEKRGSLFPMPMETKIVNP 628  
 DB 496 ESMADADIDVPGKIGINSNSLVTLSLSNRREGSPDRGGLS 580  
 QY 629 GTSVTS---WLPSPYQ---MPSPALISLHPVVAS 662  
 DB 549 AVSLAANPMPMPAPAPAAHVPMPPEAI--AHLPVFAAWTD 587

RESULT 3  
 Q88914 PRELIMINARY; PRT; 555 AA.

AC Q88914;  
 DT 01-JUN-2002 (TEMBLrel. 21, Created)  
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)  
 DE Putative ovule development protein antilegumenta (ANT).  
 GN CUI175C11.21 OR CUI1017C11.11.  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Eukaryota; Magnoliopsida; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wang R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Currie J.,





RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: U41339; AAB17364.1; -  
 DR EMBL: U40256; AAA91040.1; -  
 DR EMBL: AL035709; CAB38923.1; -  
 DR EMBL: AL161592; CAB80440.1; -  
 DR EMBL: AY080706; AAB85024.1; -  
 DR EMBL: AY117207; AAM51282.1; -  
 DR PIR: S71365; S71365.  
 DR TRANSFAC: T02639; -  
 DR GO: GO:0005634; C:nucleus; IEA.  
 DR GO: GO:0003700; F:transcription factor activity; IEA.  
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 2.  
 DR PRINTS: PR00367; ETRSPLELMT.  
 DR Prodom: PD001423; TF\_ERF; 2.  
 SQ SEQUENCE 555 AA; 61725 MW; BD275D2F082B519B CRC64;

Query Match 37.9%; Score 1328; DB 10; Length 555;  
 Best Local Similarity 43.5%; Pred. No. 1.5e-87;  
 Matches 314; Conservative 69; Mismatches 113; Indels 226; Gaps 22;

QY 1 MKRINESNNTDGNHNMVLFGLSPH-MKM-----EATGATVPTTFM 43  
 DB 1 MKSFCDNDNNHNTNLLGLFSLSNMAMKMGGRGGRGREALYSSSTSSAATSSSVPLV 60  
 QY 44 SPSGSHLSNFGMCYGVGENGFHSPDLTMPLKSDGS:CIIEALKRSQTVWVPTSPKLE 103  
 DB 61 G---DNTSNFGVCGSNFNGGTYSHMSWMPURSDSLCMLAKRSHSHNHQDSFKVE 117  
 QY 104 DFLGATWGTGTHYSGHSG--LSLDSIYNSQNAEAPNRDLISQPF---RQGSHSVQ 157  
 DB 118 DFFG---THNNHTHKKAMDLSDLSLFTYTH--EBNTTNNQEFSSPQTNH---- 166  
 QY 158 THPIYSGLAGHGLYAPLEETTKETHVSDCSLMPQTEGLKMWVAPTRFSTHQVLE 217  
 DB 167 -----EET--RNYGND-----PSLTHG----- 182  
 QY 218 QGMNCGMGNRNGVLSGVGCEL-QSLSLMSPGSSQSSCVTABS----- 261  
 DB 183 -----GSFNVGV-----YGFQQSLSLMSPGSSQSSCITGSHHQQONQNHQSNH 229  
 QY 262 -----GTDVAVDAKRGH---AKLGQKQPVYRKSIDTFGQRTSQYRGVTR 304  
 DB 230 QOISEALVETSVGFETTTAAAKKRGQEDVYVVGQKQIVHRKSIDTGGRTSQYRGVTR 289  
 QY 305 HMTGTYEAHLMDNSCKKSGGTRKGRQYVLLGGYMEEKARAYDAAALKYGPSTHLNFS 364  
 DB 290 HMTGTYEAHLMDNSCKKSGGTRKGRQYVLLGGYMEEKARAYDAAALKYGPSTHLNFS 349  
 QY 365 IENYQVLEEMKMSRQEVYVAHLRKSSGFSRGAISYRGVTRHHQGMQARIGRYAGNK 424  
 DB 350 AENYQKEIKEDMKMTKQEVYVAHLRKSSGFSRGAISYRGVTRHHQGMQARIGRYAGNK 409  
 QY 425 DLYIGFSTQFEAAEAYDVAATKFGANAVTNFDSRYDVERIMASSNLLAGE:ARRKD 484  
 DB 410 DLYIGFSTQFEAAEAYDVAATKFGANAVTNFDSRYDVERIMASSNLLAGE:ARRK-- 467  
 QY 485 NDPKNDIDYKSVTVSVNNEETVQVQAGNNNNENDESKVLLFNHPQOQOQANANGSDQ 544  
 DB 468 -----NNSIV--VENTE-----DQ 479  
 QY 545 KIMNCGYRNSAASVALDGLIGDSVSGQHMMDSSKIGTHFSNTSLVTSLSAREA 604  
 DB 480 TALNA-----VEGGSNREV 494  
 QY 605 SPEKRGPSL--LFPMPPEYTKIV-NDIGTSVTSWLPSPVQKPPPAISLSLHLYPASWT 661  
 DB 495 STPERLSPFAIPALPQVNVQKMGSMGMSPMWTSNPAELK--TVALTLTPMPVFAAMA 553  
 QY 662 DT 663  
 DB 554 DS 555

RESULT 5

Q8S1E1 PRELIMINARY; PRT; 692 AA.  
 ID Q8S1E1;  
 AC Q8S1E1;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Putative oule development protein aintegumenta-like protein.  
 GN P0035F12.3.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0035F12.3";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP003313; BAB8946.1; -  
 DR Gramene: Q8S1E1; -  
 DR GO: GO:0005634; C:nucleus; IEA.  
 DR GO: GO:0003700; F:transcription factor activity; IEA.  
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 2.  
 DR PRINTS: PR00367; ETRSPLELMT.  
 DR Prodom: PD001423; TF\_ERF; 2.  
 DR SWART: SM00380; AP2; 2.  
 SQ SEQUENCE 692 AA; 71515 MW; 4D5A0B49ED8712AF CRC64;

Query Match 27.3%; Score 956; DB 10; Length 692;  
 Best Local Similarity 35.3%; Pred. No. 1.7e-60;  
 Matches 263; Conservative 100; Mismatches 229; Indels 152; Gaps 28;

QY 16 HNWLGFSLSPHMKMEXT-----SAATVPTTFYSPGSHLSNFGMCYGVGENGFHSP 68  
 DB 5 NNMIAFSLSPDOLPSPQSTNSTLISAAATTT-----AGDSSTDVCENIPQDW----- 53  
 QY 69 LTMPLKSDGS:CIIEALKRSQTVWVPTSPKLEDFLGATMG--THEYGS----- 119  
 DB 54 -----SMGSELSALV--AEKLEDFLGISFSGQGHGKGVIPES 95  
 QY 120 -----ERGLSDSIYNSQNAEAPNRDL--SOPFQGMHVSQYTHPIYSGLAGHGL 170  
 DB 96 AACAASGSSSVQVLYPPSSSSSLQFADSVWVATSP-----VVAH---DVSQGM 144  
 QY 171 YQAPLEETTKETHVSDCSLMPQTEGLKMWVAPTRFSTHQVLEQQM--CGMGENB 229  
 DB 145 VSAASAAAAAANGGIG-----LSMKNMWR--SQAPQAPQALSLSMNAAGTTAGG 194  
 QY 230 GVSLGSV-GGCEL-----QSLSLMSPGSSQSSCVTA-----PSGDS--VAVDAYK 272  
 DB 195 GGMALLAGGERGRTTPPASESLSTSAHGAATTATMAGKKEINEEGSGAGAVVAAGSES 254  
 QY 273 RGH---AKLGQKQPVYRKSIDTFGQRTSQYRGVTRHMTGTYEAHLMDNSCKKSGGTRG 329  
 DB 255 GSGGAVVEAGAAAAAAAKRSYDVTGQRTSIYRGVTRKRWGRYEAHLMDNSCKKSGGTRG 314  
 QY 330 RQVYLLGGYMEEKARAYDAAALKYGPSTHLNFSIENYQVLEEMKMSRQEVYVAHLR 369  
 DB 315 RQ---GGYDEKEKARAYDAAALKYGPSTTTTFPVNNEYKELEEMKHTRGCFVASTLR 371  
 QY 390 KSGGFSRGAISYRGVTRHHQGMQARIGRYAGNKDLYIGTSTQREAAEAYDVAIKR 449  
 DB 372 KSGGFSRGAISYRGVTRHHQGMQARIGRYAGNKDLYIGTSTQREAAEAYDVAIKR 431  
 QY 450 GANAVTNFDSRYDVERIMASSNLLAGE:ARRKNDPRNK-DIDYKSVTVSVNNEETV 508

```

Db      432  GLAAVTVNPFMSRVDVYSILDSAAALPVGTAAKRLKDAEAAAADYVG---RIASHLGGGDAY 468
Oy      509  QVQAQNNNNENDESEKMKVLFN-----HP-----SQQQQ-----ANGNG 541
Db      489  AARHGHHSAAAAMPTIAFQAAAADPPHAGLYHPYAQPLRGWCKKODHAIVTAAASHL 548
Oy      542  SDQKIMNCG--NYRNSAFSVALODLIGDSV-----GSGHNNL-----DESSKIGTHFS 589
Db      549  QDLHLNLGAAAAAHDFPSGAMQGGGSLSDNASLHSTGNSVYVNGDNGGGGGGYIM 608
Oy      590  NTSSLVTSLSSSRASEPKRGPSSLFPMPMETKYV--NPYSTVTSMLBPPTQMPSP 644
Db      609  APMASVATATATAVAASSHDHGGDGGKQVOMGYDSYLVGADAYGGGAGMSPMAWTPASAP 666
Oy      648  AILSLH-----LPPFASMTDT 663
Db      669  AATSSDMTGVCAGCAQLFSYVNDT 692

```

## RESULT

```

AD QBLSN1 PRELIMINARY; PRT; 584 AA.
TC QBLSN1
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE BABY BOOM.
GN BMM.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C24;
RA Boutilier K., Offringa R., Sharma V.K., Kieft H., van Lammeren A.A.M.,
RA Ouellet T., Zhang L., Hatoori J., Liu C.-M., Miki B.L.A.,
RA Custers J.B.M., van Lookeren Campagne M.M.;
RT "Optic expression of the Brassica napus BABY BOOM gene triggers a
RT conversion from vegetative to embryonic growth."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
EMBL: AF317907; AACMS3803.1; -.
DR GO; GO:0005654; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 2.
DR PRINTS; PR00367; ETRSPLEMT.
DR Prodom; PD001423; TF_ERF; 2.
DR SMART; SM00380; AP2_2.
SQ SEQUENCE 584 AA; 64515 MW; 1EA3CDD1C900FB8 CRC64;

```

Query Match	27.2%;	Score 953;	DB 10;	Length 584;
Best Local Similarity	36.2%;	Pred. No. 2.2e-60;		
Matches 260;	Conservative 75;	Mismatches 174;	Indels 210;	Gaps 24

```

QY      16 HNNWLGPSLSPMKMKREATSAATVPTTYYMBSQSHLSNFMCVCGVEGNENHPSPLTYWPLK 75
      |||:|||||
Db      5 NNNWLGFSLSIPR-----DQNH-----HRTVDVSTTRTAVD 34
      |||:|||||
QY      76 SDGSLCI-LELRKRSOTQVWVWPTSSP--KLEDFLGATMTGHEYSHERGLSDSITYN 131
      |||:|||||
Db      35 VAGGYCPDLAAPSDESSAVQTSFLSPGVTLEAF-----TRDNHSHSRMDI----- 81
      |||:|||||
QY      132 SQNBAEPNRDLISQPFROGSHSVQTHPYSGSLACHGLYQAPLEERTKETHYSDCSL 191
      |||:|||||
Db      82 -----NGACNNTL-----TNN----- 93
      |||:|||||
QY      132 MPQMTGLKQNVAPTRF-FSTHQGVLEQQNCCMGERNVGLG-----SYCGC 239
      |||:|||||
Db      94 --QNGPKLENTLGGRTTITTYTNNETVVDNGDCGGGSGGSSLSLGMTKYLSSHSHVANA 151
      |||:|||||

```

```

OY      240 ELQ-----SLSLSMSGSSCSVTPASPSTSDVAVADAKKRGHAKGQCPHRRKSIDTFGQ 234
Db      152 NHODNNGARGKJSLSNKNSSTSDNNNNYNNDDV---VQEKTIVDVVEFTP--KXTIESFGQ 206

OY      295 RTSQYRGVTRHARTGYEAHLMDNSCKKEGOTRKGROYLYGGYIMEKKAAPAYDLAALKY 354
Db      207 RTSIYGVGVTRHRTGTGYEAHLMDNSCKKEGOTRKGROYLYGGYIKREKARAAYDLAALKY 266

OY      355 WGSTRTNFINSIENYQVLEEMKMGKROEYVAHLRRKSGSGFRGASIRGVYTRHHORMQ 414
Db      267 WGSTTTTNPPLSEYEKEVEBEMKMTROEYVASLRKSGSGFRGASIRGVYTRHHORMQ 326

OY      415 ARIGRVAGNDLYLJFTSTOEBAEAAYDVAIKFRGANAVTNPDISRIVERIMASSNL 474
Db      327 ARIGRVAGNDLYLJFTGTOEBAEAAYDAIKFRGLSAVTNPFMNRYNVAALIESPLP 386

OY      475 AGELARRKKNDRPNRNDIYNKSVVTSVNNETVQVAGN--NNNENDSEWKVYLFHNP 532
Db      387 TGSARLKD-----YNNPVPAMMISNNVSEBANVSGMONTAFHQ 429

OY      533 -----QQQOANGSDOKIMNCGYRNASAFMALDYLIGIDSVSGOHNMDESSKIG 585
Db      430 GMDLSILQQQKERYVG---YNNGGLSTESTRVCFK---OBEEQCHFLRNSPSHMT 479

OY      586 T--HFSNTSLSLVSLSSREASPEKRGSLLEFPMPMETKTVNPLGTSVWMLSPFV-- 641
Db      480 NYDHHSSTDSDVTVAGNVASVGGYQGFAT-----PVGISV--NYDPFTLAEI 523

OY      642 -----QMRSP-----AISLSH-----LTFVFASTDT 663
Db      526 AYNAARHNYYYAHQHQQQQQQQQSPGGGFVPAIINNHSNNVYFHGGEGGEGAFTFVWMDT 584

```

## RESULT

```

ID      08LSN2.      PRELIMINARY;      PRT;      579 AA.
AC      08LSN2.
AD      01-OCT-2002 (TReMBuRel. 22, Created)
DT      01-OCT-2002 (TReMBuRel. 22, Last sequence update)
DT      01-JUN-2003 (TReMBuRel. 24, Last annotation update)
DE      AP2/EREBP transcription factor BABY BOOM2.
GN      BM2.
OS      Brassica napus (Rape).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophytes; Magnoliophyta, eudicotyledons, core eudicots; rosids;
OC      eurosids II; Brassicales; Brassicaceae; Brassica.
OX      NCBI_TaxID=3708;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=cv. Topas.
RA      Boutilier K., Offringa R., Sharma V.K., Kieff. H., van Lammeren A.A.M.,
RA      Zhang L., Hactori U., Onellet T., Liu C.-M., Miki B.L.A.,
RA      Custers J.B.M., van Lookeren Campagne M.M.;
RT      "Ecotypic expression of the Brassica napus BABY BOOM gene triggers a
RT      conversion from vegetative to embryonic growth.";
RL      Submitted (NOV-2000) to the EMBL/GenBank/DDA databases.
EMBL: AF317905; C:ncucleus; IFA.
DR      GO; GO:0005634; C:nucleus; IFA.
DR      GO; GO:0003700; F:transcription factor activity; IFA.
DR      GO; GO:006355; P:regulation of transcription, DNA-dependent; IFA.
DR      InterPro: IPR001471; TF_ERF.
DR      Pfam: PF00847; AP2-domain; 2.
DR      PRINTS: PR00367; ETRHSPLEMT.
DR      PRODOM: PD001423; TF_ERF; 2.
DR      SMART: SM00360; AP2; 2.
SQ      SEQUENCE      579 AA; 64018 MW; 642ED3B2B06D3BEA CRC64;

```

Query Match	27.0%;	Score 947;	DB 10;	length 579;
Best Local Similarity	39.1%;	Pred. No. 5.9e-60;		
Matches 243; Conservative	65;	Mismatches 155;	Indels 158;	Gaps 22

```

15 NNNWLGESLSP-----HMKMEATGATVPT-----TFYMSPSQSHLSNFGMCYGV 59
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

DB 2 NNNWLGSLSPYEQNHHRKDVCSSTTTTAVDAGECYDPTAASDESSAIGTSPSPBGV 61  
 QY 60 -----GENGNHSPPLTWPLKSDGSLCTLEALKKSQVQVWVPVSSPKLEPDLGATMGT 114  
 DB 62 VLDAFTDRNNSSH-----RDWDINGSAAC--NNIHND-----ODGPKLENFLGRTT----- 105  
 QY 115 EYGSHERGLSDSIYNSQNAEAQPNRDLISQPFQOGHMSVQTHPYSGIACHGLYOAP 174  
 DB 106 -----TINYNTENV-----GDI-----DSSGCVG----- 124  
 QY 175 LBEETKETHVSDCSLMPQWTEGLKXNVAPTRFEFTHQOYLEQOMCMGNERNGVSLG 234  
 DB 125 -----GDDGGGSLGLSM--IKTW-----LRNPVDVNDQENG--NG 158  
 QY 235 SVCCGELQSLISMSPGSQSSCVTAPSGTDSVAVDAKKGHAKLGQKQPVHRKSIDTFQ 294  
 DB 159 AKG-----LSLSMNSTSCDNNNSNNVLAQKTIIDS-----VEATPKKTIESTFG 206  
 QY 235 RTSQYRGVTRHRTGRYEAHLMDNSCKEKGQTRKGRQVYLGQYMEKAARAAYDLAALKY 354  
 DB 207 RTSIYRGVTRHRTGRYEAHLMDNSCKEKGQTRKGRQVYLGQYMEKAARAAYDLAALKY 266  
 QY 355 WGPSTHINFSIENYQVLEEMKMSROEYVAHLRKKSGFSRGASITRGVTRHGHGRWQ 414  
 DB 267 WGTITTFPMSSEYEKEIEEMKMTROEYVASLRKSSGFSRGASITRGVTRHGHGRWQ 326  
 QY 415 ARIGVAGNKDLYLGTFTOEBAEAAYDVAIKFRGANAVTFDISRYDVERIMASSNL 474  
 DB 327 ARIGVAGNKDLYLGTFTOEBAEAAYDVAIKFRGLAVTFDMNRVYVAILDESPSLP 386  
 QY 475 AGELARRKDDPRKNDIDYKSVTSYNNMETVQVQAGNNNNDEEMKAVLEFNHS-- 532  
 DB 387 IGSAKRKE--ARRPVSMWMTISNNVSESE-----NNASGMQAAVQHQGV 432  
 QY 533 -----OQOQANGSGDOKIMNCNYRN--SAFMALODLIGISVSGQHMLDESSKIGT 586  
 DB 433 DLSLLQHQERYNG--YYNGNLSSESARACFQOE-----DDQHFL----- 473  
 QY 587 HFSNTSLVTSLSSEASPE 607  
 DB 474 --SNTQSLMTNIDHQSSVSD 492

RESULT 8  
 DB 081303 PRELIMINARY; PRT; 579 AA.  
 AC 081303;  
 DT 01-OCT-2002 (Tremblrel. 22, Created)  
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE AP2/ERBP transcription factor BABY BOOM.  
 GN BEM1.  
 OS Brassica napus (Rape).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC euroside II; Brassicales; Brassicaceae; Brassica.  
 OC NCB1\_Taxid=3708;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Topas.  
 RA Bouillier K., Offringa R., Sharma V.K., Kieft H., van Lammeren A.A.M.,  
 RA Queller T., Zhang L., Hattori J., Liu C.-W., Miki B.L.A.,  
 RA Custers J.B.M., van Lookeren Campagne M.M.,  
 RT "Ecotypic expression of the Brassica napus BABY BOOM gene triggers a  
 RT conversion from vegetative to embryonic growth."  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF317904; AA033800.1;  
 DR EMBL; AF317906; AA033802.1;  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 2.

DR PRINTS; PR00367; ETRHSPLELMENT.  
 DR Prodom; PD001423; TF\_ERF; 2.  
 DR SMART; SM00380; AP2\_2.  
 DR SEQUENCE 579 AA; 63962 MW; EA2AF24734500D72 CRC64;  
 Query Match 26.9%; Score 942; DB 10; Length 579;  
 Best Local Similarity 38.9%; Pred. No. 1.4e-59;  
 Matches 243; Conservative 61; Mismatches 156; Indels 164; Gaps 21;  
 QY 15 NNNWLGSLSP-----HMKCATSAATVPT-----TFYMSPSQSLSNFMQCYGV 59  
 DB 2 NNNWLGSLSPYEQNHHRKDVCSSTTTTAVDAGECYDPTAASDESSAIGTSPSPBGV 61  
 QY 60 -----GENGNHSPPLTWPLKSDGSLCTLEALKKSQVQVWVPVSSPKLEPDLGATMGT 114  
 DB 62 VLDAFTDRNNSSH-----RDWDINGSAACNNIHND-----ODGPKLENFLGRTT----- 105  
 QY 115 EYGSHERGLSDSIYNSQNAEAQPNRDLISQPFQOGHMSVQTHPYSGIACHGLYOAP 174  
 DB 106 -----TINYNTENV-----GDI-----DSSGCVG----- 124  
 QY 175 LBEETKETHVSDCSLMPQWTEGLKXNVAPTRFEFTHQOYLEQOMCMGNERNGVSLG 234  
 DB 125 -----GDDGGGSLGLSM--IKTW-----LRNPVDVNDQENG--NG 158  
 QY 235 SVCCGELQSLISMSPGSQSSCVTAPSGTDSVAVDAKKGHAKLGQKQPVHRKSIDTFQ 294  
 DB 159 AKG-----LSLSMNSTSCDNNNSNNVLAQKTIIDS-----VEATPKKTIESTFG 206  
 QY 235 RTSQYRGVTRHRTGRYEAHLMDNSCKEKGQTRKGRQVYLGQYMEKAARAAYDLAALKY 354  
 DB 207 RTSIYRGVTRHRTGRYEAHLMDNSCKEKGQTRKGRQVYLGQYMEKAARAAYDLAALKY 266  
 QY 355 WGPSTHINFSIENYQVLEEMKMSROEYVAHLRKKSGFSRGASITRGVTRHGHGRWQ 414  
 DB 267 WGTITTFPMSSEYEKEIEEMKMTROEYVASLRKSSGFSRGASITRGVTRHGHGRWQ 326  
 QY 415 ARIGVAGNKDLYLGTFTOEBAEAAYDVAIKFRGANAVTFDISRYDVERIMASSNL 474  
 DB 327 ARIGVAGNKDLYLGTFTOEBAEAAYDVAIKFRGLAVTFDMNRVYVAILDESPSLP 386  
 QY 475 AGELARRKDDPRKNDIDYKSVTSYNNMETVQVQAGNNNNDEEMKAVLEFNHS-- 532  
 DB 387 IGSAKRKE--ARRPVSMWMTISNNVSESE-----NNASGMQAAVQHQGV 432  
 QY 533 -----OQOQANGSGDOKIMNCNYRN--SAFMALODLIGISVSGQHMLDESSKIGT 586  
 DB 433 DLSLLQHQERYNG--YYNGNLSSESARACFQOE-----DDQHFL----- 473  
 QY 587 HFSNTSLVTSLSSEASPE 607  
 DB 474 --SNTQSLMTNIDHQSSVSD 492

RESULT 9  
 DB 09CAH3 PRELIMINARY; PRT; 425 AA.  
 AC 09CAH3;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Putative AP2 domain transcription factor.  
 GN F28P22.24.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC euroside II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCB1\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia.  
 RX MEDLINE=2106719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,

RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao O., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Elgu P., Feldblum T.V., Feng J.D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hizat L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurta D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,  
 RA Miltschev J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,  
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utecherback T., Van Aken S., Vaynsberg M., Vysotskaya V.S., Walker M.,  
 RA Yu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,  
 RA "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 thaliana.", 816-820(2000).  
 RL Nature 408:816-820(2000).  
 DR EMBL, AC010926; AAG51860.1; -.  
 DR PIR, B96750; B96750.  
 DR GO, GO:0005634; C:nucleus; IEA.  
 DR GO, GO:0003700; F:transcription factor activity; IEA.  
 DR GO, GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro, IPR001471; TF\_ERF.  
 DR Pfam, PF00847; AP2-domain; 2.  
 DR PRINTS, PR00367; ETRSPLEMT.  
 DR PRODOM, PD001423; TF\_ERF; 2.  
 DR SMART, SM00380; AP2; 2.  
 SQ SEQUENCE 425 AA; 48205 MW; C191BA52D6C5ACBB CRC64;

Query Match 26.4%; Score 923.5; DB 10; Length 425;  
 Best Local Similarity 43.4%; Pred. No. 1.9e-58;  
 Matches 214; Conservative 60; Mismatches 108; Indels 111; Gaps 16;

QY 18 WLGLSLSPHMKM-----EATSAATVPTTFYMSPSGSHLSNFGMCYGVGENGNFSLTV 71  
 DB 4 WLGFLTPPRKCNSEEBELHNDGDLTF-----THRPVWR 42  
 QY 72 MFLKSDGLCTLEALKSOTQVWPTSSPKLEDFLGATMGTHGYGHERGLSDSYNN 131  
 DB 43 YDINH-----HHHEDV-----PKVEDLLS-----NSHTEYPTN--HN 76  
 QY 132 SQNAEAQPNRDLISQPRQGHMSVQTHPYSGLACHGLYQAPLEETKETHYSDCSL 191  
 DB 77 QTNVACTTVVKNLP-----GYL-----LHDQVVTPEHY--PNT 109  
 QY 192 MPQMTG-----LKNWV--APTRFSTHQVLEQNMCGM--GNERNVSL 233  
 DB 110 DPNLSNDYGFERVSVSFVKSMLEQTPAPFLSHYVTEAGTSMNISHPSNETGYNT 169  
 QY 234 -GSVGGELQSLSLSPSGSSCYTAPAGT--DSVAVDK-KRGAHKLQKOPVARK 287  
 DB 170 NSGM-----LSLASHGACSDLINSSVSAKVEPKVDEKRLVYVQVYESVPRK 222  
 QY 288 SIDTFQORTSQYRGVTRRWTRRYEALHWDNSCKEGQTRKGRQVYLLGQYDMEKAAAY 347  
 DB 223 SYDSYQRTSQYRGVTRRWTRRYEALHWDNSCKEGQTRKGRQVYLLGQYDMEKAAAY 282  
 QY 348 DLAAALKYMPSTHINSIENYOYLEEMKMSROBYVHLBRKSGSFRGASIVYGRH 407  
 DB 283 DLAAALKYMPSTHINSIENYOYLEEMKMSROBYVHLBRKSGSFRGASIVYGRH 342  
 QY 408 HOGHGMQARIGRVAGNKDYLGTFSQEEBAAYVAALKFRGANAVTNPDISRYVERI 467  
 DB 343 HOGHGMQARIGRVAGNKDYLGTFSQEEBAAYVAALKFRGANAVTNPDINRYVERI 402  
 QY 468 MASSNLAGELAR 480  
 DB 403 CSSTIVDSDAQ 415

ID Q8LQ03 PRELIMINARY; PRT; 597 AA.  
 AC Q8LQ03;  
 DT 01-OCT-2002 (TRENBLREL. 22, Created)  
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)  
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)  
 DE Oule development aintegumenta-like protein BNM3.  
 GN BNM3.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrharioideae; Oryzaceae; Oryza.  
 OX NCBI\_Taxid=4530;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Bi X.-Z.;  
 RT "Cloning and identification of two ovule development proteins,  
 RT aintegumenta-like protein in rice (Oryza sativa)."  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AY062180; MAL47205.1; -.  
 DR Gramene, Q8LQ03; -.  
 DR GO, GO:0005634; C:nucleus; IEA.  
 DR GO, GO:0003700; F:transcription factor activity; IEA.  
 DR GO, GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro, IPR001471; TF\_ERF.  
 DR Pfam, PF00847; AP2-domain; 2.  
 DR PRINTS, PR00367; ETRSPLEMT.  
 DR PRODOM, PD001423; TF\_ERF; 2.  
 DR SMART, SM00380; AP2; 2.  
 SQ SEQUENCE 597 AA; 62198 MW; F856BEC99ADE25B CRC64;

Query Match 26.3%; Score 919.5; DB 10; Length 597;  
 Best Local Similarity 36.0%; Pred. No. 6.1e-58;  
 Matches 253; Conservative 79; Mismatches 206; Indels 165; Gaps 24;

QY 16 HNWGLSLSPHMKM-----EATSAATVPTTFYMSPSGSHLSNFGMCYGVGENGNFSLTV 75  
 DB 5 NWMLAPSLSPDO-----LPPSQT-----NSTFISAAATTTTA 37  
 QY 76 SD---SLCTLEALKSOTQVWPTSSPKLEDFLGATMGTHGYGHERGLSDSYNN 132  
 DB 38 GDSSTDVCF-----NIPQAHF-----STPAIGNGGIGLSMIKWLSS 75  
 QY 133 QNAEAQPNRDLISQPRQGHMSVQTHPYSGLACHGL--YQAPLEETKETHYSDCS 190  
 DB 76 QPA-PQP-AQPLSLSMNMGATTAAQGGAMALLAGERGRTTPASELSSTAHHGATTP 133  
 QY 191 LMPQMTGELKNWVAPTRFSTHQVLEQNMCGMERNRGSVLSGSGELQSLSLSP 250  
 DB 134 ---MAGRKET-----NEBSSGSAQAV-----VAV 155  
 QY 251 GSGSSCYTAPAGTDSVAVDKRGHAKLQKOPVARKSIDTFQORTSQYRGVTRRWTR 310  
 DB 156 GSE-----SGSGSAVVEA-----GAAAAAAKSVDTFGQRTSIRGVTRRWTR 200  
 QY 311 YEALHWDNSCKEGQTRKGRQVYLLGQYDMEKAAAYDLAAALKYMPSTHINSIENYOY 370  
 DB 201 YEALHWDNSCKEGQTRKGRQ---GYDDEKAAARAYDLAAALKYMPSTHINSIENYOY 257  
 QY 371 QLEEMKMSROBYVHLBRKSGSFRGASIVYGRVTRHHOGRWQARIGRVAGNKDYLGT 430  
 DB 258 ELEMKMTQEFVVASIRKSGSFRGASIVYGRVTRHHOGRWQARIGRVAGNKDYLGT 317  
 QY 431 FSTQEEBAAYVAALKFRGANAVTNPDISRYVERIYVASSNLAGELARRKQNDPRNK 490  
 DB 318 FSTQEEBAAYVAALKFRGANAVTNPDISRYVERIYVASSNLAGELARRKQNDPRNK 377  
 QY 491 -DIDYKSVTVSYNNEETVQVQAGNNNENDESMKNTLVN-----HP----- 531  
 DB 378 YDVG---RASHGAGGGAAYAAHYGHHHSAAMPTIAFOAAAPPPHAGLYHPYAQPL 434  
 QY 532 ---SQOQO-----ANGSGDQKINMG--NYNSAFSMALQDLIGDSV-----GSG 573  
 DB 435 RGWCKQEDHAAVIAAASHLQDLHLHLNLGAAAAAAHDFFSQMQOQHGLGSDNALSHTG 494

QY 574 QHNL---DESKITGTHFNSWISLVYSLSSREASPEKGPILLFPMEMETKIV--NPI 628  
 Db 495 SNSVYVNGDNGGGGGYIMAPMSAVSATATAVASSHDGGGGKQVOMGYDYLVDGADY 554  
 QY 629 GTSVTSMLEPTVQMRPSPALISLH-----LPVFASWTDT 663  
 Db 555 GGGAGNRMPFWMTMPASAPATISSDMTGYCHGAQLFSVWNTD 597

RESULT 11  
 QY6D4 PRELIMINARY; PRT; 655 AA.  
 ID QY6D4; (Tremblrel. 25, Created)  
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE OSUNBA0014K14.16 protein (OSUNB0116K07.8 protein).  
 GN OSUNBA0014K14.16 OR OSUNB0116K07.8.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Euphorbiaceae; Oryzae; Oryza.  
 CX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,  
 RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,  
 RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,  
 RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,  
 RA Huang Y.C., Li Y., Zhu Q.J., Zhao Q., Hu X., Liu Y.L., Mu Z.,  
 RA Chen L., Fan D.L., Meng J.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,  
 RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,  
 RA Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao F., Zhang L., Wu M.,  
 RA Zhang R.Q., Guan J.P., Hong G.F.,  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,  
 RA Liu Y.L., Mu Z., Yu Z., Chen L., Fan D.L., Wang Q.J., Zhang L.,  
 RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,  
 RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen X.Y.,  
 RA Hao F., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,  
 RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,  
 RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,  
 RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,  
 RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.,  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL606604; CAB02944.1;  
 DR EMBL; AL731638; CAB05555.1;  
 SQ SEQUENCE 655 AA; 67291 NM; DEFCCD07729A6933 CRC64;

Query Match 26.2%; Score 918.5; DB 10; Length 655;  
 Best local similarity 42.8%; Pred. No. 8.2e-58;  
 Matches 227; Conservative 51; Mismatches 117; Indels 135; Gaps 18;

QY 17 NWLGFSLT-----PHMTEATSAATVPTTFYMSPSOSHISNFCMGVGENGNFSPITLV 71  
 Db 6 NWLGFSLTGGCNPHQNGSPSA-----GDAIDISGSDPFGULTPPAHITGA- 56  
 QY 72 MPLKSDGSLCTLEALKR-----SOTCWMPYS----- 98  
 Db 57 ---GEDAPYGVDAFNGGTHETDPMANGLDYGGSSDLSMLVSGSGGRTVAADGYE 113  
 QY 99 SPKLEDFLGATM-----GTHEY-----GSHERG-----LSLDSIYNSON 134  
 Db 114 APRLENFLDGNFSDVHGAAGVLYSGAVGAGVNSGGCGGTIELSMIKTLRNSQ 173  
 QY 135 AEAQPNRLDLSQPRQGHNSVQTHPIYSGLAHGILYQAPLEBETTKETHVSDCSLMPQ 194  
 Db 174 SQOOP-----SPF---QAADQGMSTDAASSYAC-----SDVLVSGCG 209  
 QY 195 MTEGLKTVAPTRFETHQVLEQNMNCGKNGERNNGVSLGVSCEGLQSLSTNSPQSGS 254

Db 210 ---GAGGTA-----SSHQGLALSTST-----SVAAAGG--CAVVAASSSEN 251  
 QY 255 SCYAPSGTDSVAVDKKRGAHAKQOKOPVHRKSIDTFQORTSOYRGVTRRWGRYEAH 314  
 Db 252 KRVDSPGG---AVDG-----AVPKSIDTFGQRTSIYRGVTRRWGRYEAH 295  
 QY 315 LMDNSCKKXGQTRKSGQVYLGVDMEKRAARVYDALIKYNGSPRHINFSIENVOVLEE 374  
 Db 296 LMDNSCKRSGQSKGR---GGYDKEDKARARVYDALIKYNGCTTTTNPMSNEKLEEE 352  
 QY 375 MKMNSRQEVYALHRRKSGFSRGAISYRGVTRHGHGWQARIGFVAGNKDLYLGTEGTQ 434  
 Db 353 MKMTRQEVYALHRRNSGFSRGAISKYRGVTRHGHGWQARIGFVAGNKDLYLGTEFSTE 412  
 QY 435 EEAARVDAVATKFRGANAVTNPDISRYDVERIMSSNLGELARRKKD 484  
 Db 413 EEAARVDAVATKFRGANAVTNPDISRYDVERIMSSNLGELARRKKD 462

RESULT 12  
 QY6F42 PRELIMINARY; PRT; 581 AA.  
 ID QY6F42; (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Ovine development protein aIntegument-like protein.  
 GN 110B6.90.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.  
 CX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Naamizu E., Kotani H.,  
 RA Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.,  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Fu Arabidopsis sequencing project,  
 RL EMBL; AL91142; CAC01738.1;  
 DR PIR; T51580; T51580.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; P:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription; DNA-dependent; IEA.  
 DR InterPro; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 2.  
 DR PRINTS; PRO0367; ETHRSPELEMT.  
 DR PRODOM; PD001423; TF\_ERF; 2.  
 DR SMART; SM00380; AP2; 2.  
 SQ SEQUENCE 581 AA; 64170 NM; 292A8ACCC29B79D CRC64;

Query Match 36.2%; Score 916.5; DB 10; Length 581;  
 Best local similarity 25.2%; Pred. No. 9.6e-58;  
 Matches 255; Conservative 76; Mismatches 170; Indels 223; Gaps 25;

QY 16 NWLGFSLTSPHNMTEATSAATVPTTFYMSPSOSHISNFCMGVGENGNFSPITLVPLK 75  
 Db 5 NWLGFSLTSPH-----DONH-----HRTVDSTTKTAVD 34  
 QY 76 SDGSLCT-LEALKRSQVYVPTSP---KLEDFLGATMGTHEYSHER-----GLSLD 126  
 Db 35 VAGCYCPDLAASDSSSAVQTSFSLPFGVTLEAF-----TRNNNSGRWMD-NGAGCN 87  
 QY 127 SIYYSQNAEAQPNRLDLSQPRQGHNSVQTHPIYSGLAHGILYQAPLEBETTKETHVS 186  
 Db 88 NINNEQON----- 96  
 QY 187 DCSLMPQMTBGLKTVAPTRFETHQVLEQNMNCGKNGERNNGVSLG----- 234



Db 97 -----BK-----LENFGRITTTIYNTEIVDNGDGGGGGGGSLGSLNKTWLSNH 146  
 Qy 235 SVGGGELQ-----SLSSMSPGSSCVTAPSGTDSVAVDAKKRGHAKLQKQPHRKS 289  
 Db 147 SVANAHHQDONGNARGSLSSMSSTSDSNNNNNNDV---VQEKTVDVETTP--KXTI 201  
 Qy 290 DTEGRTSQRGTGRHNTGRYEAHLMDNSCKEGQTRKGRQVYLGTYMEKARAYDL 349  
 Db 202 ESFGKRTSTYGRVTRHMTGRYEAHLMDNSCKEGQTRKGRQ---GGYDKKARAYDL 258  
 Qy 350 AALKYMGSTRHNSIENYQVLEEMKMSRCEYVAHLRKSGSGFSRGSITRYGTRHQ 409  
 Db 259 AALKYMGTTTTTNPPLSEYKEVEKMTROGYVASLKRSGSGFSRGSITRYGTRHQ 318  
 Qy 410 HGRVQATIGRYAKNDLYLTSTQEEAAAYVAALKRGANATNPDISYDVERIMA 469  
 Db 319 HGRVQATIGRYAKNDLYLTSTQEEAAAYVAALKRGASAVTNPMKRYNKAILE 378  
 Qy 470 SSNLAGELARRKKNDPRNKDIDYKSVYVSNNETVQVQAGN--NNNENDEMKAVL 527  
 Db 379 SPSP.LIGSSAKRLKD-----VNNVPMAMLSNNVSEANNVSGMONTA 421  
 Qy 528 FNPSP-----QQQANGSGSOKIMNGCNTRNSAFSALODLIGDSVSGQNMDE 580  
 Db 422 FOHHQMDLSLQQQOERYVG---YNGGNLSTESTRVCFK---QEEBQCHFLNNS 471  
 Qy 581 SSKIGT--HFSNTSLVTSLSRSREASPEKRGPSLLFPMPMETKIVNPIGTSVSLPS 638  
 Db 472 PSHMTNVDSHSSSDSVTCGVVSGYGGQFAL-----PVGTSV--NDPFF 517  
 Qy 639 PTV-----QMRSP-----AISLSH-----LPVPS 659  
 Db 518 TAAEIAVNAHNYTYAHOQQOQOQOQSGPQDPFVALSNHSMNTFHGGGEGAGAPTSV 577  
 Qy 660 WTD 663  
 Db 578 WNDT 561

RESULT 13  
 Q9F1E2 PRELIMINARY; PRT; 555 AA.  
 AC Q9F1E2;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Similarity to unknown protein.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.  
 OC NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Columbia;  
 RK MEDLINE=9156233; Pubmed=10048468;  
 RA Asamizu E., Sato S., Kaneko T., Nakamura Y., Kobani H., Miyajima N.,  
 RA Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.  
 RT Sequence features of the regions of 1,081,956 bp covered by seventeen  
 RT physically assigned pl and YAC clones.";  
 RL DNA Res 5:379-391(1998).  
 DR EMBL, AB016891; BAB08476.1;  
 DR GO:GO:0005634; C:nucleus; IEA.  
 DR GO:GO:0003700; F:transcription factor activity; IEA.  
 DR GO:GO:0006355; P:regulation of transcription, DNA-dependency; IEA.  
 DR InterPro: IPR001471; TF\_SRF.  
 DR Pfam: PF00847; AP2-domain; 2.  
 DR PRINTS: PR00367; ETRHSPLEMT.  
 DR Prodom: PD001423; TF\_SRF; 2.  
 DR SMART: SM00380; AP2\_2.  
 QO SEQUENCE 555 AA; 59930 MW; 881FL43370E3A3A1 CRC64;

Query Match 25.7%; Score 900.5; DB 10; Length 555;  
 Best Local Similarity 36.0%; Pred. No. 1.3e-56;  
 Matches 245; Conservative 78; Mismatches 182; Indels 175; Gaps 22;  
 Qy 7 SNNTDGNNHNNLSPSLS---PHMKMTSAATVPTTMYMSPSGHNLFPKCVGANG 63  
 Db 19 SPSSSSSSHHNNLSTSLNNNNNNFSSNPULSTSDHHHPHSLF----- 68  
 Qy 64 NFHSPLTPML-KSDGSLCILALRSQTYV-VPTSSPKEDFLGATWGTGHEGHER 121  
 Db 69 ---QAFSTSPVRODSPGVSPS---DATAVLSTVPCGPKLENFPGCASSTT----- 114  
 Qy 122 GLSLSDIYNNQNAQAQPRDLISQPFQGGHMSYQTHPPYSGLAChGLYQAPLEETTK 181  
 Db 115 -----TTPMQO-----VQS---LGGVFSDDLPLPSPSA 144  
 Qy 182 ETHVSDCSLMPQMTGLKNVVAPTREFSTHQVLEQMGNGMNGRNGVSLGSGCEL 241  
 Db 145 EIVDSLKSLIASF-----LGVSGHSEVSVAKQ 176  
 Qy 242 QSLSSMSPGSSCVTAPSGTDSVAVDAKKRGHAKLQKQPHRKSITDEGRTSQYRG 301  
 Db 177 QRPPLVSEAS-----P7PKQVSEFGQSTISYRG 206  
 Qy 302 VTRHNTGRYEAHLMDNSCKEGQTRKGRQVYLGTYMEKARAYDLAALKYMGPSYTH 361  
 Db 207 VTRHNTGRYEAHLMDNSCKEGQTRKGRQ---GGYDKKARAYDLAALKYMGPTTT 263  
 Qy 362 NFSIENYQVLEEMKMSRCEYVAHLRKSGSGFSRGSITRYGTRHQHGRVQATIGRYA 421  
 Db 264 NPFISYSELEEMKMTROEFVSLRKSGFSRGSITRYGTRHQHGRVQATIGRYA 323  
 Qy 422 GNKDLYLTSTQEEAAAYVAALKRGANATNPDISYDVERIMASNL-LAGELAR 480  
 Db 324 GNKDLYLTSTQEEAAAYVAALKRGANATNPDISYDVERIMASNL-LAGELAR 382  
 Qy 481 RKNDPRNKDIDYKSVYVSNNETVQVQAGN--NNNENDEMKAVL 525  
 Db 383 PSPATADKTYDLSPSDSLTTSPSLTFNVAITPVNDGGTFYHTGDIKDPADHYWSN 442  
 Qy 526 VLFNHSQOQANGSGSOKIMNGCNTRNSAFSALODL-----GIDSVSGQHNM 577  
 Db 443 IF-----GFQANPRAEKRPILANFGSDLHNSPGYAIWVQGEENNFGSFGVSDGNN 496  
 Qy 578 LBESSKIGHFNTSLVTSLSRSREASPEKRGPSLLFPMPMETKIVNPIGTSVSLPS 637  
 Db 497 HSBAS-----NPVSAIPUSSTTTWSNGNEGIGN-----INMINNIS-- 535  
 Qy 638 SPTVQMRSPPAISLHPVF 657  
 Db 536 --SYCTAKS-NISVLTPTVF 552

RESULT 14  
 Q9LT37 PRELIMINARY; PRT; 540 AA.  
 AC Q9LT37;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Similarity to unknown protein.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.  
 OC NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Columbia;  
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;  
 RL Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.  
 QO SEQUENCE FROM N.A.

RC STRAIN=Columbia;  
 RX MEDLINE=20277480; PubMed=10819329;  
 RA Nakamura Y;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
 RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC  
 RT clones.";  
 RL DNA Res. 7:131-135(2000).  
 DR EMBL; AB025629; BAB02492.1; -;  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; P:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR01471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 2.  
 DR PRINTS; PR00367; ETRKSPLEMT.  
 DR ProDom; PD001423; TF\_ERF; 2.  
 DR SMART; SM00380; AP2; 2.  
 SQ SEQUENCE 540 AA; 59066 MW; 0A013E0735176DDD CRC64;

Query Match 25.5%; Score 893.5; DB 10; Length 540;  
 Best Local Similarity 38.2%; Pred. No. 4e-56;  
 Matches 238; Conservative 74; Mismatches 142; Indels 169; Gaps 23;

DB 100 PKLEDFLGATGCTGHEGSHERGSLDSTIYNSQNAEQPNKDLISQFFRQGHMSVQTH 159  
 27 PKVADPLG--VSKPDENQSNHLVAYVDSDFYFHTNS-LMP-----SVQSN 68  
 QY 160 PYYSLAGHGLVQAPLEETTKETHYSDSSLMPTGSLKQWVAFTREFTHQVLEQQ 219  
 DB 69 -----DVVVAACDSVTPNN-----SSYHLEQSA 92  
 QY 220 MNGCMGNERNVSLGSGGELQSLSLMSPGSSCVTAPSGTDSVAVDKAKRHAYLG 279  
 DB 93 HN-----LQSLTSMGTTAGNNVVDKAPSETTGDNAGCALAYVE 133  
 QY 280 QKQPVHRKSIDTGCQTSQYRGVTRHRTGRYEAHLMNSCKKEQTKRGQVYLYGYDM 339  
 DB 134 TATP--RRALDTPGQRTSIRGVTRHRTGRYEAHLMNSCKKEQTKRGQVYLYGYDM 188  
 QY 340 EEEKARAYDLAALKYGPSTHINFSIENTYQVLEEMKMSRQEVYVAHLRKSQSGSRGAS 399  
 DB 189 EDKAAISYDLAALKYGPSTHINFSIENTYQVLEEMKMSRQEVYVAHLRKSQSGSRGAS 248  
 QY 400 IYGVTRHOGHGMQAFGRVANKDLYIGTSTQEEAAEAVDAIKRGANATNDI 459  
 DB 249 MYGVTRHOGHGMQAFGRVANKDLYIGTSTQEEAAEAVDAIKRGANATNDI 308  
 QY 460 SRVDPRIMASSNL-LAGE-LARKD-----NDPRNKID----- 493  
 DB 309 NRVDKALLESSTLPIGGGAARLKEAQLSESRKEAEMLAGSFQYGGSGTSGSGT 368  
 QY 494 -----YKSY-----VTSVNNBETVQVQANNNDSEMKVLPNHSQ----- 533  
 DB 369 SSRLOQPLPPLSTIQPLEPLSLQND--ISHYNNNNADSS--SFNHSYTOTQHA 421  
 QY 534 -----QOANGNSDQKIMNCGYNSAFSMALODLIGDSVGSQAHMLDESSXI 584  
 DB 422 LHQCTNNYLOOQSSQ--SQOLNAYIHSNPL--LHGIVSTSIYDNNNN-----GSSS 472  
 QY 585 GTRFSNTSLVTSI-----SSSRPASEPKRGPSLLFPMPEMTKIYVPIGTSVTSWLPSP 639  
 DB 473 GSY--NTAALFQNHGIGIGSSSTVGSTEE-----PPTVTDYDMSPSDGTGSGWTSE 524  
 QY 640 TYQMRPSPALSLSLPVEASMTD 662  
 DB 525 SVQ-GSNPG-----GVFTYMN 540

RESULT 15  
 ID Q41832 PRELIMINARY; PRT; 485 AA.  
 AC Q41832;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Of protein.  
 GN ORF.  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.  
 CC NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Black Mexican Sweet, TISSUE=Endosperm;  
 RA Daniel T.J., Fordham-Skelton A.P., Vergani P., Edwards R.;  
 RT "Isolation of a Maize cDNA (Accession No. Z47554) Encoding AETA1A2-  
 RT Like Binding Domains by Complementation Cloning of an L-Isoaspartyl  
 RT Methyltransferase-Deficient Mutant of Escherichia coli (PGR 96-013).";  
 RL Plant Physiol. 110:1435-1435(1996).  
 DR EMBL; Z47554; CAA87634.1; -;  
 DR PIR; T03638; T03638.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; P:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 2.  
 DR ProDom; PD001423; TF\_ERF; 2.  
 DR SMART; SM00380; AP2; 2.  
 SQ SEQUENCE 485 AA; 52391 MW; 2B165E3D17FC8B3A CRC64;

Query Match 25.4%; Score 890; DB 10; Length 485;  
 Best Local Similarity 44.2%; Pred. No. 6.1e-56;  
 Matches 200; Conservative 45; Mismatches 94; Indels 114; Gaps 10;

QY 187 DCSLMPQMTGKKNVAPFRETSTQVLEQCMNG--MGNERKGV-----LG 234  
 DB 4 DMSAYPH-----HMLSFSLNNYHGLLEAFNSGCTPLGDEQAVEBSPFTVEDFLG 57  
 QY 235 SVG-----CGELQSLSLMSPGSSQS-----CVTAPSGT 263  
 DB 58 GVGCVGAPRRRLQIRITSLCAASCSITRFLRHPAQSCTVGPESLFLAAMST 117  
 QY 264 DSAVADAKKRGHAKLGQKQPVHRKSIDTGCQTSQYRGVTRHRTGRYEAHLMNSCKE 323  
 DB 118 DVAAEASDQAS-----RSAETFGQRTSIRGVTRHRTGRYEAHLMNSCKE 165  
 QY 324 GQTFKRGVYLVGVDMEKARAYDLAALKYGPSTHINFSIENTYQVLEEMKMSRQEV 383  
 DB 166 GQSKRGQVYLVGVDMEKARAYDLAALKYGPSTHINFSIENTYQVLEEMKMSRQEV 225  
 QY 384 VAHLRKSQSGSRGASIRGVTRHOGHGMQAFGRVANKDLYIGTSTQEEAAEAYDV 443  
 DB 226 IASLRKKSQSGSRGASIRGVTRHOGHGMQAFGRVANKDLYIGTSTQEEAAEAYDI 285  
 QY 444 AALFRGANVTPNDISRYVERIM-----ASSNLLAG 476  
 DB 286 AALFRGANVTPNDISRYVERIM-----ASSNLLAG 345  
 QY 477 ELARKKNDPNRNDI-----DYNSVVTYNNBETVQVQANN 515  
 DB 346 AAASQATMPSEKDYVSLALAHYQOOQOEPRQPAASAYAYGSGGVNVDFTMTSGSN 405  
 QY 516 NNENDS-EW-----KMTLFNHPSOQOQANGGSD 543  
 DB 406 NNTSGVYMGATSGAVVGOQDSSKQNGYASN 438

Search completed: March 9, 2004, 10:47:03  
 Job time : 49 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2004, 10:29:58 ; Search time 60 Seconds  
(without alignments)  
3122.150 Million cell updates/sec

Title: US-10-024-632-2

Perfect score: 3502  
Sequence: 1 MKRINESNNITDDGNHNLG.....RPSPAISLSHPFASMTDT 663

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3502	100.0	663	5	ABR79636 Soybean A
2	1715	49.0	665	5	ABR79637 Soybean A
3	1408.5	40.2	642	5	ABR79638 Rice AINT
4	1349.5	38.5	669	5	ABR79639 Rice AINT
5	1338	37.9	555	3	ABR07724 An Arabid
6	1328	37.9	555	3	AA31380 Arabidops
7	1328	37.9	555	3	AA31380 Arabidops
8	1328	37.9	555	3	ABR40832 Arabidops
9	1328	37.9	555	3	ABR40830 Arabidops
10	1295.5	37.0	548	3	AA31382 Arabidops
11	1295.5	37.0	548	3	AA31381 Arabidops
12	1180.5	33.7	585	5	ABR79640 Cotton AI
13	1061.5	30.3	557	6	ABR40856 Glycine m
14	1059.5	29.9	710	6	ABR40768 Glycine m
15	1046.5	29.9	707	6	ABR40768 Glycine m
16	1021.5	29.2	384	6	ABR40770 Glycine m
17	956	27.3	632	6	ABR40877 Oryza sat
18	948.5	27.1	510	6	ABR40764 Glycine m
19	948	27.1	574	7	ADD30879 Plant yie
20	948	27.1	574	7	ADD30834 Plant yie
21	948	26.9	574	7	ADD30834 Plant yie
22	942	26.9	579	4	ABR50873 Brassica
23	942	26.9	579	4	ABR50872 Brassica
24	939.5	26.8	415	7	ADD30560 Plant yie
25	936	26.7	553	6	ABR40771 Glycine m

25	934	26.7	538	6	ABR40854 Goseypium
27	931.5	26.6	470	6	ABR40853 Catalpa s
28	927	26.5	486	6	ABR40767 Glycine m
29	927	26.5	558	7	ADD30981 Plant yie
30	923.5	26.4	425	6	ABR40833 Arabidops
31	922.5	26.3	489	6	ABR40756 Zea mays
32	922.5	26.3	512	6	ABR40852 Glycine m
33	920.5	26.3	530	6	ABR40769 Glycine m
34	918	26.2	469	6	ABR40759 Oryza sat
35	911.5	26.0	484	6	ABR40757 Zea mays
36	893.5	25.5	540	6	ABR40837 Arabidops
37	891.5	25.5	498	7	ADD30350 Plant yie
38	890	25.4	485	6	ABR40829 Zea mays
39	890	25.4	485	6	ABR40834 Zea mays
40	879	25.1	516	6	ABR40836 Arabidops
41	874	25.0	492	6	ABR40855 Zea mays
42	874	25.0	566	6	ABR40876 Arabidops
43	831	23.7	160	6	ABR40761 Glycine m
44	717	20.5	312	6	ABR40762 Glycine m
45	687	19.6	412	6	ABR40851 Glycine m

## ALIGNMENTS

RESULT 1  
ABR79636  
ID ABR79636 standard; protein; 663 AA.  
AC ABR79636;  
XX  
XX 21-OCT-2002 (first entry)  
XX  
XX Soybean AINTEGUMENTA-like polypeptide GmANT1.  
XX  
XX AINTEGUMENTA; ANT-like polypeptide; GmANT1; soybean; transgenic plant;  
XX plant.  
XX  
XX Glycine max.  
XX  
XX  
XX WO200259332-A2.  
XX  
XX 01-AUG-2002.  
XX  
XX 19-DEC-2001; 2001MO-US049294.  
XX  
XX 21-DEC-2000; 2000US-0257896P.  
XX  
XX (MONS ) MONSANTO TECHNOLOGY LLC.  
XX  
XX He SS, Doteon SB;  
XX  
XX WPI; 2002-559798/64.  
XX  
XX N-PSDB; ABRN84480.  
XX  
XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in  
PT improving agronomic, horticultural, and quality traits of plants, such as  
PT increased size of plant organs.  
XX  
XX Claim 3; Page 130-133; 165pp; English.  
XX  
XX The present sequence is the protein sequence of GmANT1, a newly  
CC identified AINTEGUMENTA-like (ANT-like) polypeptide of soybean. The  
CC sequence is predicted from clone CPR67663 (see ABRN84480), which was  
CC obtained from a soybean sequence database screening using Arabidopsis ANT  
CC as query. Also isolated was clone CPR67626 encoding GmANT2 (see  
CC ABR79637). GmANT1 and GmANT2 show homology to ANT in 2 N-terminal AP2 DNA  
CC binding domains, but have C-terminal sequences that bear little, if any,  
CC homology to ANT although they share conserved segments with each other.  
CC The invention provides nucleic acid encoding ANT-like polypeptides  
CC comprising, in the N-terminal to C-terminal direction, 2 AP2 DNA binding  
CC domains followed by an amino acid subsequence selected from those given  
CC in ABR79629-35. ANT-like polypeptides have been identified in soybean.

rice, cotton and corn (see ABB79636-41). Nucleic acids (see ABBN84480-86) encoding the ANT-like polypeptides can be used in the construction of transgenic plants, especially corn, soybean, canola, wheat, cotton, tomato or potato (all claimed) having improved agronomic, horticultural or other quality traits, such as increased size of plant organs. These plants are especially useful for production of ethanol or animal feedstuff.

Sequence 663 AA;

Query Match 100.0%; Score 3502; DB 5; Length 663;

Best Local Similarity 100.0%; Pred. No. 3, 1e-316; Indels 0; Gaps 0; Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MKRINESNNTDDGNHNMILGFSLSPHMKMEATSAATVPTTFYMSBSQSHLSNFGMCYGVG 60
DB 1 MKRINESNNTDDGNHNMILGFSLSPHMKMEATSAATVPTTFYMSBSQSHLSNFGMCYGVG 60
QY 61 ENGNTFHSPLTWMPKSDGSLCTLEALKRSQTVQVWPSSPKLEDFLGATWGTHEYSGHE 120
DB 61 ENGNTFHSPLTWMPKSDGSLCTLEALKRSQTVQVWPSSPKLEDFLGATWGTHEYSGHE 120
QY 121 RGLSLDSIYNSQNAEAPNRDLISQPRQGHMSVQTHPYSGLAGHLYQAPLEERTT 180
DB 121 RGLSLDSIYNSQNAEAPNRDLISQPRQGHMSVQTHPYSGLAGHLYQAPLEERTT 180
QY 181 KETHYSDCSLMPQMTGKXNVAPTRFSTHQVLEQNMNGMNEENGSLSGVCGE 240
DB 181 KETHYSDCSLMPQMTGKXNVAPTRFSTHQVLEQNMNGMNEENGSLSGVCGE 240
QY 181 KETHYSDCSLMPQMTGKXNVAPTRFSTHQVLEQNMNGMNEENGSLSGVCGE 240
DB 181 KETHYSDCSLMPQMTGKXNVAPTRFSTHQVLEQNMNGMNEENGSLSGVCGE 240
QY 241 LQSLISMSPGSSCCTAPSGTDSVAVDAKKRGHAKLQKQPVHRKSIDTFGQRTSQYR 300
DB 241 LQSLISMSPGSSCCTAPSGTDSVAVDAKKRGHAKLQKQPVHRKSIDTFGQRTSQYR 300
QY 301 GYTRHRTGRYEAHLMDNSCKEGRKQRYVLCGYMEERAAAYLLAALKTWGPSTH 360
DB 301 GYTRHRTGRYEAHLMDNSCKEGRKQRYVLCGYMEERAAAYLLAALKTWGPSTH 360
QY 361 INPSIENYQVLEEMKNGMSROEVVAHLRKSQSGFRGASITYGVRHHQGRWQARIQY 420
DB 361 INPSIENYQVLEEMKNGMSROEVVAHLRKSQSGFRGASITYGVRHHQGRWQARIQY 420
QY 421 AGNKDYLIGTSTOEBAEAAYVAALKEFGANAVTNFDISRYDVERIVASSNLLAGELAR 480
DB 421 AGNKDYLIGTSTOEBAEAAYVAALKEFGANAVTNFDISRYDVERIVASSNLLAGELAR 480
QY 481 RKXNDPRKXNDIDYNSVYTSVYNNERTVOVQAGNNNNENDSEMKVYLFNHPSSQOQOQANGN 540
DB 481 RKXNDPRKXNDIDYNSVYTSVYNNERTVOVQAGNNNNENDSEMKVYLFNHPSSQOQOQANGN 540
QY 541 GSDQKIMNGYNSAFSVALQDLIGIDSVSGQHNMLEDSSKIGTHFENTSSLVTSLS 600
DB 541 GSDQKIMNGYNSAFSVALQDLIGIDSVSGQHNMLEDSSKIGTHFENTSSLVTSLS 600
QY 601 SRASAEKRGPSLLFPMPMETKIVNPICTSVTSMLPSPTVQMBSPALSLHLVPAWSM 660
DB 601 SRASAEKRGPSLLFPMPMETKIVNPICTSVTSMLPSPTVQMBSPALSLHLVPAWSM 660
QY 661 TDT 663
DB 661 TDT 663

```

RESULT 2

ABB79637 standard; protein; 665 AA.

XX ABB79637;  
XX 21-OCT-2002 (first entry)  
XX Soybean AINTEGUMENTA-like polypeptide GmANT2.  
XX

KM AINTEGUMENTA; ANT-like polypeptide; GmANT2; soybean; transgenic plant;  
KW plant.

XX Glycine max.

XX WO200259332-A2.

XX 01-AUG-2002.

XX 19-DEC-2001; 2001WO-US049294.

XX 21-DEC-2000; 2000US-0257896P.

XX (MONS) MONSANTO TECHNOLOGY LLC.

XX He SS, Doseon SB;

XX WPI; 2002-599798/64.

XX N-PSDB; ABBN84481.

XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in improving agronomic, horticultural, and quality traits of plants, such as increased size of plant organs.

XX Claim 3; Page 137-140; 169pp; English.

The present sequence is the protein sequence of GmANT2, a newly identified AINTEGUMENTA-like (ANT-like) polypeptide of soybean. The sequence is predicted from clone CBR6762 (see ABBN84481), which was obtained from a soybean sequence database screening using Arabidopsis ANT as query. Also isolated was clone CBR6763 encoding GmANT1 (see ABB79636). GmANT1 and GmANT2 show homology to ANT in 2 N-terminal AP2 DNA binding domains, but have C-terminal sequences that bear little, if any, homology to ANT although they share conserved segments with each other. The invention provides nucleic acids encoding ANT-like polypeptides comprising, in the N-terminal direction, 2 AP2 DNA binding domains followed by an amino acid subsequence selected from those given in ABB79629-35. ANT-like polypeptides have been identified in soybean rice, cotton and corn (see ABB79636-41). Nucleic acids (see ABBN84480-86) encoding the ANT-like polypeptides can be used in the construction of transgenic plants, especially corn, soybean, canola, wheat, cotton, tomato or potato (all claimed) having improved agronomic, horticultural or other quality traits, such as increased size of plant organs. These plants are especially useful for production of ethanol or animal feedstuff.

SQ Sequence 665 AA;

Query Match 49.0%; Score 1715; DB 5; Length 665;

Best Local Similarity 54.4%; Pred. No. 7, 4e-150; Matches 379; Conservative 84; Mismatches 166; Indels 68; Gaps 20;

```

QY 1 MKRINESNNTDDGNHNMILGFSLSPHN-----KMEATSAATVPTTFYMSBSQSHLSN 52
DB 1 MKRINESNNTDDGNHNMILGFSLSPHNIGVSSHQSSAAAVVPTSFHHTAP--LSS 58
QY 53 FGMCGYV-GENGNFHSPLTWMPKSDGSLCTLEALKRSQTVQVWPSSPKLEDFLGATW 111
DB 53 FGMCGYV-GENGNFHSPLTWMPKSDGSLCTLEALKRSQTVQVWPSSPKLEDFLGATW 118
QY 59 YGFIYGLAENVGLYSALPIMPLKSDGSLXGLETLSRSQQAAMATSTPLENPLGSEAM 118
DB 59 YGFIYGLAENVGLYSALPIMPLKSDGSLXGLETLSRSQQAAMATSTPLENPLGSEAM 118
QY 112 GT-HEY--GSHERGSLDSIYNSQNAEAPNRDLISQPRQ-----QGHMSVQTHPY 162
DB 112 GT-HEY--GSHERGSLDSIYNSQNAEAPNRDLISQPRQ-----QGHMSVQTHPY 162
QY 119 GTPHHECSAETETWPLSLDSVFY-IPSRDRDPNNNOTYQNHVQHISTNQOQOQELQAY 177
DB 119 GTPHHECSAETETWPLSLDSVFY-IPSRDRDPNNNOTYQNHVQHISTNQOQOQELQAY 177
QY 163 SGLACHGLYQAPLEERTKETHYSDCSLMPQ-----MTGLKXNVAPTRF-STHQ 213
DB 163 SGLACHGLYQAPLEERTKETHYSDCSLMPQ-----MTGLKXNVAPTRF-STHQ 213
QY 178 STLNDHMYL-----ESSKQSGTSDNNNLVQNMGGDAPVYGLKSM--EYVNFQASHA 230
DB 178 STLNDHMYL-----ESSKQSGTSDNNNLVQNMGGDAPVYGLKSM--EYVNFQASHA 230
QY 214 QVLEQNMCGMGERNGV--SLGSGGGLQSLSLMSPGSSCCTV-----APSGTDSY 266
DB 214 QVLEQNMCGMGERNGV--SLGSGGGLQSLSLMSPGSSCCTV-----APSGTDSY 266
QY 231 H--ESKVIIVHVENAGESGSGISGMAYGDIQSLSLMSPGSSSVTSYSHRASPAVVDV 288
DB 231 H--ESKVIIVHVENAGESGSGISGMAYGDIQSLSLMSPGSSSVTSYSHRASPAVVDV 288
QY 267 AVDAKKRGHAKLQKQPVHRKSIDTFGQRTSQYRGVTRHRTGRYEAHLMDNSCKEGRGT 326

```

DB 269 AMDTKKKEPEVDDKQIVHRSITDPGQRISQYGVTRHMTGTYEHLMDNCKEKGOS 348  
QY 327 RKGEVYLYGYDMEKKARAYDIALKXWGPSTHINSINYQVLEEMQNGROEVVAH 386  
DB 349 RKGEVYLYGYDMEKKARAYDIALKXWGPSTHINPLINYQVLEEMQNGROEVVAH 408  
QY 387 LRKSSGSGSRGASLYRGVTRHOGRWQARIGRVAGKNDLYLGFPTQEEAAAYDVAI 446  
DB 409 LRKSSGSGSRGASLYRGVTRHOGRWQARIGRVAGKNDLYLGFPTQEEAAAYDVAI 468  
QY 447 KFRGANAVTNFDIRYDVERIMASSNLAGELARKKNDPRNKDIDYKNSVTSV-NNE 505  
DB 469 KFRGANAVTNFDIRYDVERIMASSNLAGELARKKNDPRNKDIDYKNSVTSV-NNE 528  
QY 506 ETVOVQAGNNNNNDSEMKVLFNHPGQQOQANGSDOKIMNCGNYRNSAFSVALODLI 565  
DB 529 EAILMHQKSCSEND-QKRWVLY-QSSQQLQONPTIE-----SDKTNSFAVALDNMF 580  
QY 566 GIDSVGSGQNNMLDESSKIGTHFNTSLSVTSLSSSSEASPEKRGPSLLFPMPMETKIV 625  
DB 581 -----HQVEESSKARTHVSNPSLSATSLSSSREGSPDRSLPMLSGMPSATSKL 631  
QY 626 NPIGTSVTSNLPSPVQKRPSPALISLTPVFAAWTD 662  
DB 632 ATNPNNVNSMDPSPLR---PALTLPQMPVFAAWTD 664

RESULT 3  
ABB79638  
ID ABB79638 standard; protein; 642 AA.  
XX ABB79638;  
AC 21-OCT-2002 (first entry)  
XX  
DT  
XX  
DE Rice AINTEGUMENTA-like polypeptide OsANT1.  
XX  
XX AINTEGUMENTA, ANT-like polypeptide; OsANT1, rice, transgenic plant;  
KM plant.  
XX  
OS Oryza sativa.  
OS  
PN WO200259332-A2.  
XX  
PF 01-AUG-2002.  
XX  
XX 19-DEC-2001; 2001WO-US049294.  
XX  
XX 21-DEC-2000; 2000US-0257896P.  
PR  
XX (MONS ) MONSANTO TECHNOLOGY LLC.  
PA  
XX He SS, Dotson SB;  
PI  
XX  
XX MPI: 2002-599798/64.  
DR N-PSDB; ABB79638.  
XX  
XX  
PT New nucleic acid encoding AINTEGUMENTA-like polypeptides useful in  
PT improving agronomic, horticultural, and quality traits of plants, such as  
PT increased size of plant organs.  
XX  
PS Claim 3; Page 143-146; 16pp; English.  
XX  
CC The present sequence is the protein sequence of OsANT1, a newly  
CC identified AINTEGUMENTA-like (ANT-like) polypeptide of rice. The sequence  
CC was predicted from a cDNA clone (see ABB79638) obtained from particle RNA  
CC by PCR amplification. OsANT1 cDNA was not amplified from root or leaf  
CC RNAs, suggesting tissue-specific expression. Initial OsANT1 clones had  
CC been identified in a rice database screening using the newly identified  
CC soybean ANT-like polypeptide GmANT1 as query. An additional rice ANT-like  
CC polypeptide, OsANT2 (see ABB79639), was also isolated. OsANT1 polypeptide  
CC shares high homology with Arabidopsis ANT, GmANT1 and GmANT2 at the N-

CC terminal AP2 DNA binding domains, shares conserved segments at the N-  
CC terminus, and shares conserved segments with GmANT1 and GmANT2, but not  
CC with ANT, at the C-terminus. The invention provides nucleic acids  
CC encoding ANT-like polypeptides comprising, in the N-terminal to C-  
CC terminal direction, 2 AP2 DNA binding domains followed by an amino acid  
CC subsequence selected from those given in ABB79629-35. ANT-like  
CC polypeptides were identified in soybean, rice, cotton and corn (see  
CC ABB79638-41). Nucleic acids (see ABB79638-41) encoding the ANT-like  
CC polypeptides can be used in the construction of transgenic plants,  
CC especially corn, soybean, canola, wheat, cotton, tomato or potato (all  
CC claimed) having improved agronomic, horticultural or plant quality  
CC traits, such as increased size of plant organs. These plants are  
CC especially useful for production of ethanol or animal feedstuff

Sequence 642 AA;  
SQ  
Query Match 40.2%; Score 1408.5; DB 5; Length 642;  
Best Local Similarity 46.5%; Pred. No. 2.4e-121;  
Matches 341; Conservative 64; Mismatches 149; Indels 179; Gaps 27;  
QY 13 GNNHMLGFSLEPHKMEATSAATVPTT-----F 41  
DB 5 GGSNWLGFSLSPHMPAMEVPSSEPTAAHHHHHPAAAAAAGMSPPDSATTCNF 64  
QY 42 YMSPGSHL-----SNFGMCYGVEN--GNFHSPLTVMPLKSDPSLCITLAKRSQTQ 92  
DB 65 LPSPPAQQVAVAPSGYTYVGAIGDTSTAGVYSHLPVMPKSDGLCTIMG----- 117  
QY 93 VNVPTSSPFLDEDFLGATWGTGHEYSGHERGLSDSIYNSQAEAPNPDLLSOPFRQG 152  
DB 118 -WMPSSPFLDEDFLGATWGTGHEYSGHERGLSDSIYNSQAEAPNPDLLSOPFRQG 160  
QY 153 HNSVQTHPIYSGLAGHLYQAPLEBETTYETHTVSDSSLMPTWTEGLKRWVAPTEFSH 212  
DB 161 HOLV---PYN-----YQPLEAMLOEAAAAPMEDAM-----AAAKNF--- 195  
QY 213 QCVLEQQNMGNGNERNVSLGSGGEL-QSLSLMSMSPSQ--SSCVT-APSGTDSVAVD 269  
DB 196 ---LVTSYACVGNQ-----EMPOSLSLMSMSPSQSSCVSAAPQHQOMAVV 240  
QY 270 A-----KRGAAKLQKQPVHKKSLDTGQRISQYGVTRHMTGTYEHLMDNCKEKGOS 307  
DB 241 AAAAAAGDQGSNSNDGEGORVGRKGTGQKQPVHKKSLDTGQRISQYGVTRHMTGTYEHLMDNCKEKGOS 300  
QY 308 TGRYEALHMDNCKEKGOTRKGQVYLYGYDMEKKARAYDIALKXWGPSTHINSINY 367  
DB 301 TGRYEALHMDNCKEKGOTRKGQVYLYGYDMEKKARAYDIALKXWGPSTHINSINY 360  
QY 368 YQVLEEMQNGROEVVAHLLRRKSSGFSRGASLYRGVTRHOGRWQARIGRVAGKNDLY 427  
DB 361 YRDEIEEMERMTQOEYVAHLRRRSSGFSRGASLYRGVTRHOGRWQARIGRVAGKNDLY 420  
QY 428 LGFTSQEEAAAYDVAIKFRGANAVTNFDIRYDVERIMASSNLAGELARKK----- 483  
DB 421 LGFTSQEEAAAYDVAIKFRGANAVTNFDIRYDVERIMASSNLAGELARKK----- 480  
QY 484 --DNDRPNKDIDYKNSVTSVNNNEETVOVQAGNNNNNDSEMKVLFNHPGQQOQAA 537  
DB 481 APDHFVIGSEL-----GATEEASAAATVGT-----TDWRMV--HGSCQQQAAACT 523  
QY 538 NGNGSQOKIMNCGNYRNSAFSVALODLIGIDSVGSGQNNMLDESSKIGTHFNTSLSV 596  
DB 524 EATADLQK-----GFWGDHLS-ALHGIIVPDVESAADSDIVDGGKISGINSNSSLVT 577  
QY 597 SLSSSEASPEKRGPSLLFPMPMETKINP-IGTSVTS---MLPSPFTVO---MRDSPA 649  
DB 578 SLSSSEASPEKRGPSLLFPMPMETKINP-IGTSVTS---MLPSPFTVO---MRDSPA 630  
QY 650 SLSHLPVFAAWTD 662  
DB 631 --AHLVFAAWTD 641

## RESULT 4

ABB79639 standard; protein; 669 AA.

ID ABB79639

AC ABB79639

DT 21-OCT-2002 (first entry)

DE Rice AINTEGUMENTA-like polypeptide OsANT2.

KW AINTEGUMENTA; ANT-like polypeptide; OsANT2; rice; transgenic plant;

XX plant.

XX Oryza sativa.

XX WO200259332-A2.

XX 01-AUG-2002.

XX 19-DEC-2001; 2001WO-US049294.

XX 21-DEC-2000; 2000US-0257896P.

XX (MONS ) MONSANTO TECHNOLOGY LLC.

XX He SS, Dotson SB;

XX WPI; 2002-599798/64.

XX N-PSDB; ABBN84483, ABBN84484.

XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in

XX improving agronomic, horticultural, and quality traits of plants, such as

XX increased size of plant organs.

XX Claim 3; Page 154-157; 169pp; English.

XX The present sequence is the protein sequence of OsANT2, a newly

XX identified AINTEGUMENTA-like (ANT-like) polypeptide of rice. The sequence

XX was predicted from a genomic DNA sequence (see ABBN84483) obtained from a

XX rice database screening. An additional rice ANT-like polypeptide, OsANT1

XX (see ABB79638), was also identified. The invention provides nucleic acids

XX encoding ANT-like polypeptides comprising, in the N-terminal to C-

XX terminal direction, 2 AP2 DNA binding domains followed by an amino acid

XX subsequence selected from those given in ABB79623-35. ANT-like

XX polypeptides were identified in soybean, rice, cotton and corn (see

XX ABB79636-41). Nucleic acids (see ABBN84480-86) encoding the ANT-like

XX polypeptides can be used in the construction of transgenic plants,

XX especially corn, soybean, canola, wheat, cotton, tomato or potato (all

XX claimed) having improved agronomic, horticultural or other quality

XX traits, such as increased size of plant organs. These plants are

XX especially useful for production of ethanol or animal feedstuff

XX

XX

XX Sequence 669 AA:

XX Query Match 38.5%; Score 1349.5; DB 5; Length 669;

XX Best Local Similarity 45.3%; Pred. No. 8e-116;

XX Matches 341; Conservative 66; Mismatches 142; Indels 203; Gaps 34;

XX 13 GNNHNMVGFSLSPHMKREAT-----SAATVPTTFYNSP 45

XX 18 GGVGWGVGFSLSPHMAVYCAGVDDVGHNNHHHHVHQOQHGGLFYNPAAVAASSFYGG 77

XX 46 SQ-----SHLSNFGVCVGVNGENFHSPLTVPLKSDGSLCILEALK--PSOTQVMVPTSS 99

XX 78 GHDAVVTSAAGGSYTAG-----FSMPKPSGSLCIMALGSGOEGQGVVVAS 129

XX 100 PKLEDFLG-GATNGTHEYSGHERGLSD--SIYNSQNAEAGCNRLDLSQPRQGHM-- 154

XX 130 PKLEDFLGAGPAM-----ALSLDNSAFYGGHG--HHQGHAD 165

XX 155 --SVQTHPYYSG--LACHGLYQAPLEERTKETHVSDCSLMPQWTEGAKWVAPTRRF 209

XX 166 GGAVGADPHHGGGSLQC-----AVIPGAGG----- 192

QY 210 STHQVLEQDNMC-----GMGNERNVSLGSVGC-----GETQSLSLSM- 248

DB 193 --HDAALVHDSAAVAAGMAAHGCGYDINAAADVDY-CAAGPIIPFGHHPITLSMS 249

QY 249 SPQSQSCVY--APSGTDSVAVDA--KKGHA-KLGQKQPYHRKSITTFQRTSQYGV 302

DB 250 SAGSQSCVYQAAAAGEPYMAMDVSKRGADACQKQPVHRKSITTFQRTSQYGV 309

QY 303 TTRWTGRYEAHLNNSCKEGQTKRGQVYIGVDMEEKARAYDLAALKYGPSTHN 362

DB 310 TTRWTGRYEAHLNNSCKEGQTKRGQVYIGVDMEEKARAYDLAALKYGPSTHN 369

QY 363 FSIENYQVLEEMKNSRQEVYAHLRKSGFSRGSASIRGVYTRHGHGRWQARI 422

DB 370 FLEDYQVLEEMKNSRQEVYAHLRKSGFSRGSASIRGVYTRHGHGRWQARI 429

QY 423 NKDLVIGTF-----STOEAEAYDVAATKRGANAVTNPDISRYVERIM 468

DB 430 NKDLVIGTFIASAFAPARARHAGTOEAEAYDVAATKRGANAVTNPDITRYDVDTKL 489

QY 469 ASENLAGELARRK-KDNPRNKDIDYKSVTVSYNNEETVOQAQNNNNENDSEMKNYL 527

DB 490 ESESTLPGSLARRKKGVDGCG-----AAAVADAAALVQAG-----NVAEMKAT 535

QY 528 ----FNHPSQOQOANGNSDQKINCGNYRNASFMALQDLIG-IDVSGSQHMLDES 581

DB 536 AALPMAARTEDOOCHGSHGHC--HHDLFPDASV-LQDIVSTVDAAGAPR----- 585

QY 582 SKIGTFHSNTSLVTSLSRSREASPEKRG-----SLFPMPMETKIVNPIGTSV 632

DB 586 --APMMSMA--TSLGNSREOSPD-RGVGGGCGGVATLPAKPAASKLYSPV--PL 636

QY 633 TSWL-PSPTVQMPSPA-ISLSHLPVFASWTD 662

DB 637 NTWASPSPAVSVPARAGVSIHLPMFAAWTD 668

## RESULT 5

AAB07724 standard; protein; 555 AA.

ID AAB07724

AC AAB07724

DT 07-NOV-2000 (first entry)

DE An Arabidopsis aintegumenta (ANT) polypeptide.

KW Aintegumenta; ANT; cell proliferation; growth; organ mass; fertility;

KW asexual reproduction; plant; male sterile plant; female sterile plant;

KW early flowering.

XX Arabidopsis sp.

XX Key Location/Qualifiers

XX Domain 281..357

XX Domain /note="AP2 domain"

XX Domain 383..451

XX Domain /note="AP2 domain"

XX WO200040694-A2.

XX 13-JUL-2000.

XX 07-JAN-2000; 2000WO-US000465.

XX 08-JAN-1999; 99US-00227421.

XX (REGC ) UNIV CALIFORNIA.

XX Fischer RL, Mizukami Y;

XX WPI; 2000-465969/40.

DR N-PSDB; AAA59220.  
 XX Modulating growth and cell proliferation in a plant used to alter organ  
 PT mass, control fertility and enhance asexual reproduction in plants  
 PT comprises modulating ANT activity and selecting plants with altered cell  
 PT number.  
 XX  
 XX Claim 11, Page 41, 54pp; English.  
 XX  
 XX The present sequence represents an aintegumenta (ANT) polypeptide. The  
 CC ANT gene is expressed and functions not only in developing ovules but  
 CC also in various developing organs. Growth and cell proliferation in plant  
 CC can be modulated by modulating ANT activity. Modulation of ANT activity  
 CC is used to alter organ mass, control fertility and enhance asexual  
 CC reproduction in plants. Increased ANT activity can be used to produce  
 CC male or female sterile plants. Inhibition of ANT activity can be used to  
 CC truncate vegetative growth, resulting in early flowering  
 CC  
 XX

Sequence 555 AA;

Query Match 37.9%; Score 1328; DB 3; Length 555;

Best Local Similarity 43.5%; Pred. No. 5.8e-114; Indels 226; Gaps 22;

Matches 314; Conservative 69; Mismatches 113;

QY 1 MRRINESNTDDGNNHNLGFSLSPH-MKM-----EATSAATVPTTFM 43  
 DB 1 MASFCDNDNNHNTNLTGFSLSNMKMGREGREAIYSSSTSAATSSSVPLQTV 60  
 QY 44 SPSGSHLSNFGMCGVGENGNHSPLTMPKSDSLCTILALRSQCVNVPSSKLE 103  
 DB 61 G---DNTSNFVGCYSNPGIYSHMSVPLRSDSLCLMALNRSSHSHHSDSSKVE 117  
 QY 104 DLGAGTGTGTHYGSHERG--LSLSIYNSQNAEOPRDLISOPF---RQCGHMSVQ 157  
 DB 118 DFG---THNNTTHKEMDLSDLSLFYNTH---EENITNTQEFSPQTNNH---- 166  
 QY 158 THPYYSGLACHGLYQALEETTKETHVSDGSLMPQMTBGLKMWAVATREFTTHQVLE 217  
 DB 167 ---EET--RNYGND-----PSLTHG----- 182  
 QY 218 QQMNGMGNERGVSLGSGCEL--QSLISMSPGSSSCVTAPS----- 261  
 DB 183 ---GSFNVG---YEFQOSLSLSMSPSSCITGSHHQNQONHQSNNH 229  
 QY 262 ---GDSVAVDAKRGH---AKLGQKOPVHRKSIDTFGQRTSQRGYTR 304  
 DB 230 QQISEALVETSVGFETTWAAKKRGQEDVVVYQKQIVHRKSIDTFGQRTSQRGYTR 289  
 QY 305 HMTGRYEAHMDNCKECCGCTRKROYLLGGYDMEKAAARAYDLAAKYGSPHTNFS 364  
 DB 290 HMTGTYEHLMDNCKEKGSKRGQYLLGGYDMEKAAARAYDLAAKYGSPHTNFS 349  
 QY 365 IENYQVLEEMKMSRQEVYAHARRKSGFSRNASIYGVTRRHQHGWRQARIGRVAGNK 424  
 DB 350 AENYQKEIYDMQMRQEVYAHARRKSGFSRNASIYGVTRRHQHGWRQARIGRVAGNK 409  
 QY 425 DLYLGFSTQEEAAEAYDAALKFRGANAYTFPDISRDYERIMASNLGELARRKCD 484  
 DB 410 DLYLGFSTQEEAAEAYDAALKFRGANAYTFPDISRDYERIMASNLGELARRKCD 467  
 QY 485 NDRPKNDIDYNSVVTSVNNEETVOYQAGNNNNENDSEMKVLFNHPSSQOOOANGNGSDQ 544  
 DB 468 -----NNSIV--VANTE-----DQ 479  
 QY 545 KIMNCGNYSASFAALQDLIGDVGSGQHMLDESSKIGTHFENTSSLYTSLSSSREA 604  
 DB 480 TALNA-----VEGSSNREV 494  
 QY 605 SPEKRGPSL--LFPMPMETKIV-NPIGTSVTSWLPPTVQRPSPALSLHLPFVFSWT 661  
 DB 495 STPERLSTFPAIFALQOVNQKFGSNMGSMNPMSINPAELK--IVALLTLPMFPFAAMA 553  
 QY 662 DT 663

DB 554 DS 555  
 RESULT 6  
 AAG31380  
 ID AAG31380 standard; protein; 555 AA.  
 XX  
 XX AAG31380;  
 AC  
 XX  
 XX 17-OCT-2000 (first entry)  
 XX  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 37675.  
 DE  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KM hybridisation assay; genetic mapping; gene expression control; promoter;  
 XX termination sequence.  
 XX Arabidopsis thaliana.  
 OS  
 XX  
 XX EP1033405-A2.  
 XX  
 XX PD 06-SEP-2000.  
 XX  
 XX  
 XX 25-FEB-2000; 2000EP-00301439.  
 PF  
 XX 25-FEB-1999; 99US-0121825P.  
 XX 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 25-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 18-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139751P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144634P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 25-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148644P.  
PR 16-AUG-1999; 99US-0149368P.

PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-015263P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159320P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 14-OCT-1999; 99US-0159684P.  
PR 18-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 22-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 25-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 37.9%; Score 1328; DB 3; Length 555;  
Best Local Similarity 43.5%; Pred. No. 5.8e-114; Indels 226; Gaps 22;  
Matches 314; Conservative 6; Mismatches 113;

QY 1 MKRINSENNITDGGNNHNLGFSLSFPH-MKN-----EATSAATVPTTFYM 43  
DB 1 MKSPCDNDNNHSHVTTLLGFSLSNNMKKGGKGRALVSSSTSSAATSSSVPPOLVV 60  
QY 44 SPQSHLSNFGMGCVGVEENGNFHSPFLTWPLKSDGSLCIIEALKRSOTQVWVPTSSPKLE 103  
DB 61 G---DNTSNFGVCTGSPNPGITVSHMSVPLRSDGSLCIIEALNRSHSHHDDSSPKVE 117



```

QY 104 DFLGATMGTHYSGSHERG--LSLDSIYNSQNAEAPNDLSQPF-----ROQGHMSVQ 157
DB 118 DFRG-----THHNTSHKEMDLSDLSLFYNTTH---EPNTTNQGFSPQTRNH-----166
QY 158 THPYSGGLACHGLYQAPLEETTKETHVSDCCSLMPQMTBSLKNWVAPRFEFTHQVLE 217
DB 167 -----EERT--RNYGND-----PSLTHG-----182
QY 218 QQNMCGMERNRNGSVLSGCGGEL-QSLSLMSPGSSGSCVTAPS-----261
DB 183 -----GSFNVGV-----YGEFOQSLSMSPGSSGSCITSHHHQONQNHQSONH 229
QY 262 -----GTDVAVDAKKRGH---AKLGKQKPYHRKSIDTFGQRTSQYRGVTR 304
DB 220 QQISEALVETSVGFETTTMAAAKKRGQEDVAVVVGKQIVARKSIDTFGQRTSQYRGVTR 289
QY 305 HRMTGRYEAHLMDNSCKKEGCTRKGRQVYLGVDMEKARAYDLALKTWGSPETHNFS 364
DB 290 HRMTGRYEAHLMDNSFKKEGSHRKGROYLGVDMEKARAYDLALKTWGSPETHNFS 349
QY 365 IENYQVQLEEMKMSRQRYVAHLRRKSSGFSRGASIRYGVTRHHQGRMQRIGRYAGNK 424
DB 350 AENYQKEIEDMKMTROEYVAHLRRKSSGFSRGASIRYGVTRHHQGRMQRIGRYAGNK 409
QY 425 DLYLGTSSTOEBAAEAYVAALIKFRGANAAVTNPDISYDVERIMASSNLLAGEIARRKD 484
DB 410 DLYLGTSSTOEBAAEAYVAALIKFRGINAVTNPDITRIDYDRIMSSNTLLSGELARRN--467
QY 485 NDPNRKQIDYKSVVTSVNNSEETVQVQAGNNNNENDSEMKVLFNHPBQOQOANGNSDQ 544
DB 468 -----NNSIV--VRNTE-----DQ 479
QY 545 KIMNGCANYRNSAFSMALQDLIGDSVSGGQNHMLDESKIGTHSNISLSYTSISSGREA 604
DB 480 TALNA-----VEGSSNKEV 494
QY 605 SPEKQSHLNFQMCYGVGENGNFHSPLTNVPLKSDGSLCILELAKSQOTVWVPISPKLE 103
DB 495 STPEELLSFPALFALPDYVQNMKFGSNMGKNSPMTSNPAELK-TVALTLTQMTVFAPAMA 553
QY 662 DT 663
DB 554 DS 555

```

```

XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harwell LT,
PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,
PI Tarczy-Harlow NC,
DB WPI, 2003-201509/19.
QY Novel nucleotide fragment encoding polypeptides having receptor-like
PT protein kinase activity, caleosin-like activity, useful for altering oil
PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
XX
PS Claim 12, Page 463-465; 542pp; English.
XX The present invention describes an isolated nucleotide fragment (I)
CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
CC polypeptide (p) having receptor-like protein kinase activity, mitogen
CC activated protein (MAP)-kinase activity, lipid-1-like transcription factor
CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
CC activity, and CMC-like transcription factor activity. Also described: (1)
CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
CC oil obtained from (V). (I) or its part can be used in antisense
CC inhibition or co-suppression in a transformed plant. (III) is useful for
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
CC creating transgenic plants having altered lipid profiles. (I) can also be
CC used as a hybridisation probe. ACC00626 to ACC00686 and ABR40591 to
CC ABR40879 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 555 AA,
Query Match 37.9%; Score 1328; DB 6; Length 555;
Best Local Similarity 43.5%; Pred. No. 5,8e-114;
Matches 314; Conservative 69; Mismatches 113; Indels 226; Gaps 22;
QY 1 MRRINSENTTDDGNHNMWLGFSLSPE-MKM-----EATSAATVPTTFYM 43
DB 1 MKSFCDNDNNHNTNTNLLGFLSSNMWMMGGRGREAIYSSSTCSAATSSSVPOLVV 60
QY 44 SPEQSHLNFQMCYGVGENGNFHSPLTNVPLKSDGSLCILELAKSQOTVWVPISPKLE 103
DB 61 G--DNTSNFGYCYGNGFNPGIYSHMSVPLSDGSLCMEALNRSNHHQDSSPKYE 117
QY 104 DFLGATMGTHYSGSHERG--LSLDSIYNSQNAEAPNDLSQPF-----ROQGHMSVQ 157
DB 118 DFRG-----THHNTSHKEMDLSDLSLFYNTTH---EPNTTNQGFSPQTRNH-----166
QY 158 THPYSGGLACHGLYQAPLEETTKETHVSDCCSLMPQMTBSLKNWVAPRFEFTHQVLE 217
DB 167 -----EERT--RNYGND-----PSLTHG-----182
QY 218 QQNMCGMERNRNGSVLSGCGGEL-QSLSLMSPGSSGSCVTAPS-----261
DB 183 -----GSFNVGV-----YGEFOQSLSMSPGSSGSCITSHHHQONQNHQSONH 229
QY 262 -----GTDVAVDAKKRGH---AKLGKQKPYHRKSIDTFGQRTSQYRGVTR 304
DB 220 QQISEALVETSVGFETTTMAAAKKRGQEDVAVVVGKQIVARKSIDTFGQRTSQYRGVTR 289
QY 305 HRMTGRYEAHLMDNSCKKEGCTRKGRQVYLGVDMEKARAYDLALKTWGSPETHNFS 364
DB 290 HRMTGRYEAHLMDNSFKKEGSHRKGROYLGVDMEKARAYDLALKTWGSPETHNFS 349
QY 365 IENYQVQLEEMKMSRQRYVAHLRRKSSGFSRGASIRYGVTRHHQGRMQRIGRYAGNK 424
DB 350 AENYQKEIEDMKMTROEYVAHLRRKSSGFSRGASIRYGVTRHHQGRMQRIGRYAGNK 409
QY 425 DLYLGTSSTOEBAAEAYVAALIKFRGANAAVTNPDISYDVERIMASSNLLAGEIARRKD 484
DB 410 DLYLGTSSTOEBAAEAYVAALIKFRGINAVTNPDITRIDYDRIMSSNTLLSGELARRN--467

```



XX asexual reproduction; plant; male sterile plant; female sterile plant;  
KW early flowering.

OS Brassica sp.  
XX  
XX WO200040694-A2.  
XX  
XX 13-JUL-2000.  
XX  
XX 07-JAN-2000; 2000MO-US000465.  
PF  
XX 08-JAN-1999; 99US-00227421.  
PR  
XX (REGC ) UNIV CALIFORNIA.  
PA  
XX  
XX  
PI Flecher RL, Mizukami Y;  
XX  
XX WPI; 2000-465969/40.  
DR  
DR N-PSDB; AAS59222.

XX  
XX Modulating growth and cell proliferation in a plant used to alter organ  
PT mass, control fertility and enhance asexual reproduction in plants  
PT comprises modulating ANT activity and selecting plants with altered cell  
PT number.

XX  
XX Claim 15; Page 44; 54pp; English.

XX The present sequence represents an aintegumenta (ANT) polypeptide. The  
CC ANT gene is expressed and functions not only in developing ovules but  
CC also in various developing organs. Growth and cell proliferation in plan  
CC can be modulated by modulating ANT activity. Modulation of ANT activity  
CC is used to alter organ mass, control fertility and enhance asexual  
CC reproduction in plants. Increased ANT activity can be used to produce  
CC male or female sterile plants. Inhibition of ANT activity can be used to  
CC truncate vegetative growth, resulting in early flowering

XX Sequence 548 AA;

Query Match	37.3%;	Score 1307;	DB 3;	Length 545;
Best Local Similarity	43.6%;	Pred. No. 5,1e-12;		
Matches	309;	Conservative	77;	Mismatches 109;
			Indels	220; Gaps
Qy	9	NTDDGNHNNTLGFSLSPFN-----KQELSATVPTFTYMSQ----	SHLSNGMCT	57
Db	6	DNDSTNTNLGLSLSSNNMLKMGGGGALYSSSSSSVATS--SVPPQVLVDNDSNNGVLY		63
Qy	58	GVG-ENGNGFSPLTWPLKSDSLCILEALKRS-----QTYWVPTSSPKLEDFLGA		109
Db	64	GSNLAAREMYSQMSVYPLASDGLCLMEALNRSNNHHNSQV-----SSPKNEDF----		115
Qy	110	TMGTHEYG-SHERG-LSLDSITYNSQNAEAPNRLDISQPFROGMSVQTHPYSGLA		166
Db	116	-FGTHHNTHSHKEMALDSLDSLFYNTTTHA--PNNNTNFCRF--FSPQTNH-----		163
Qy	167	CHGLYCARLESEETTXDTHVSDCSLMPQWBSLKNWAPRSTFSTHQVLEQNGCMGN		226
Db	164	-----DEET--RNYEND-----PGATHG-----		179
Qy	227	ERNGVSLGSGCEEL--QSLSLMSPGSGSCVTA-----P		260
Db	180	---GSPFNVGVEPFOQSLSLMSPGSGSCITAHNNHQNQTNHOOISALVETSAGE		236
Qy	261	SGTDSVAVNDAKKXG-HAKXGOKOPYHRSIDTFQGRSQYGVGRHMTGRYEALHNDNS		319
Db	237	TTTMAAAAKKKGQEVVVGQOKIYHRSIDTFQGRSQYGVGRHMTGRYEALHNDNS		296
Qy	320	CKKEGQTRKRGROYLYGGDYDMEKAAAYDLAALKTWGSTHINFISINYOVLSEMKNS		379
Db	297	FKKGSHRKRGQVLYGGDYDMEKAAAYDLAALKTWGSGSTHTNFSVENYOKEIDDKMNT		356
Qy	380	ROEVVALRRKSGSFSGASTYRGVRRHHQRMQARIGRAYKNKDIYLGFESFOEAAE		439
Db	357	ROEVVAHRRKTSFSGASTYRGVRRHHQRMQARIGRAYKNKDIYLGFQGEAAE		416

QY	44	AYVVAAIKRGANAATNEDISSYDERIMASSNLLAGELARRKKDNPBKIDINYKSV	4499
Db	417	AYVVAAIKRGNAATNEDITRIYDVRIMASNTLLSGEMARRNS-----NSIV	465
QY	500	TSVNNEE--TVQVAGANNNNENDESEKMTLFLNHPSCQOQANGSGQKIMNCGRNSA	556
Db	466	RNISDEEALTLVNVGSGKREV-----GSEPRYLSE-----	496
QY	557	FSVALDDLIGIDVSGGQNMLEDSEKICGTFHSNTSLVTSLSSSRASEPEKGPSLLFP	616
Db	497	-----PRIFA	501
RESULT 10			
XX	AGG31382		
ID	AGG31382 standard; protein; 528 AA.		
XX	AC		
XX	AA	AGG31382;	
DT	17-OCT-2000	(first entry)	
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 37677.	
XX	XX		
KW	XX	Protein identification; signal transduction pathway; metabolic pathway;	
KM	XX	hybridisation assay; genetic mapping; gene expression control; promoter;	
XX	XX	termination sequence.	
OS	Arabidopsis thaliana.		
XX	XX		
XX	XX	EP1033405-A2.	
XX	PD	06-SEP-2000.	
XX	PF		
XX	25-FEB-2000;	2000EP-00301439.	
PR	25-FEB-1999;	99US-0121825P.	
PR	05-MAR-1999;	99US-0123180P.	
PR	09-MAR-1999;	99US-0123488P.	
PR	23-MAR-1999;	99US-0125788P.	
PR	25-MAR-1999;	99US-0126264P.	
PR	29-MAR-1999;	99US-0126785P.	
PR	01-APR-1999;	99US-0127462P.	
PR	06-APR-1999;	99US-0128234P.	
PR	08-APR-1999;	99US-0128714P.	
PR	16-APR-1999;	99US-0129845P.	
PR	19-APR-1999;	99US-0130077P.	
PR	21-APR-1999;	99US-0130449P.	
PR	23-APR-1999;	99US-0130510P.	
PR	23-APR-1999;	99US-0130891P.	
PR	28-APR-1999;	99US-0131449P.	
PR	30-APR-1999;	99US-0132048P.	
PR	30-APR-1999;	99US-0132407P.	
PR	04-MAY-1999;	99US-0132484P.	
PR	05-MAY-1999;	99US-0132485P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	06-MAY-1999;	99US-0132487P.	
PR	07-MAY-1999;	99US-0132863P.	
PR	11-MAY-1999;	99US-0134256P.	
PR	14-MAY-1999;	99US-0134218P.	
PR	14-MAY-1999;	99US-0134219P.	
PR	14-MAY-1999;	99US-0134221P.	
PR	14-MAY-1999;	99US-0134370P.	
PR	18-MAY-1999;	99US-0134768P.	
PR	19-MAY-1999;	99US-0134941P.	
PR	20-MAY-1999;	99US-0135124P.	
PR	21-MAY-1999;	99US-0135533P.	
PR	24-MAY-1999;	99US-0135629P.	
PR	25-MAY-1999;	99US-0136021P.	

PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140355P.  
PR 24-JUN-1999; 99US-0140356P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142300P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 26-JUL-1999; 99US-0145313P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147303P.  
PR 05-AUG-1999; 99US-0147193P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149368P.  
PR 18-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149266P.  
PR 20-AUG-1999; 99US-0149267P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 25-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 27-AUG-1999; 99US-0151083P.  
PR 31-AUG-1999; 99US-0151388P.  
PR 01-SEP-1999; 99US-0151330P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 11-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155559P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158322P.  
PR 12-OCT-1999; 99US-0158839P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160880P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.



```
PR 06-JUL-1999; 99US-0142390P
PR 08-JUL-1999; 99US-0142803P
PR 09-JUL-1999; 99US-0142920P
PR 12-JUL-1999; 99US-0142977P
PR 13-JUL-1999; 99US-0143542P
PR 14-JUL-1999; 99US-0143624P
PR 15-JUL-1999; 99US-0144005P
PR 16-JUL-1999; 99US-0144085P
PR 16-JUL-1999; 99US-0144086P
PR 19-JUL-1999; 99US-0144325P
PR 19-JUL-1999; 99US-0144331P
PR 19-JUL-1999; 99US-0144332P
PR 19-JUL-1999; 99US-0144333P
PR 19-JUL-1999; 99US-0144334P
PR 19-JUL-1999; 99US-0144335P
PR 20-JUL-1999; 99US-0144352P
PR 20-JUL-1999; 99US-0144632P
PR 20-JUL-1999; 99US-0144684P
PR 21-JUL-1999; 99US-0144814P
PR 21-JUL-1999; 99US-0145086P
PR 21-JUL-1999; 99US-0145088P
PR 22-JUL-1999; 99US-0145085P
PR 22-JUL-1999; 99US-0145087P
PR 22-JUL-1999; 99US-0145089P
PR 22-JUL-1999; 99US-0145192P
PR 23-JUL-1999; 99US-0145145P
PR 23-JUL-1999; 99US-0145218P
PR 23-JUL-1999; 99US-0145224P
PR 26-JUL-1999; 99US-0145276P
PR 26-JUL-1999; 99US-0145313P
PR 27-JUL-1999; 99US-0145918P
PR 27-JUL-1999; 99US-0145919P
PR 28-JUL-1999; 99US-0145951P
PR 02-AUG-1999; 99US-0146386P
PR 02-AUG-1999; 99US-0146388P
PR 03-AUG-1999; 99US-0146389P
PR 03-AUG-1999; 99US-0147038P
PR 04-AUG-1999; 99US-0147204P
PR 05-AUG-1999; 99US-0147302P
PR 05-AUG-1999; 99US-0147192P
PR 05-AUG-1999; 99US-0147260P
PR 06-AUG-1999; 99US-0147303P
PR 06-AUG-1999; 99US-0147416P
PR 09-AUG-1999; 99US-0147493P
PR 09-AUG-1999; 99US-0147935P
PR 10-AUG-1999; 99US-0148171P
PR 11-AUG-1999; 99US-0148319P
PR 12-AUG-1999; 99US-0148341P
PR 13-AUG-1999; 99US-0148565P
PR 13-AUG-1999; 99US-0148684P
PR 16-AUG-1999; 99US-0149368P
PR 17-AUG-1999; 99US-0149175P
PR 18-AUG-1999; 99US-0149426P
PR 20-AUG-1999; 99US-0149722P
PR 20-AUG-1999; 99US-0149723P
PR 20-AUG-1999; 99US-0149929P
PR 23-AUG-1999; 99US-0149902P
PR 23-AUG-1999; 99US-0149930P
PR 25-AUG-1999; 99US-0150566P
PR 26-AUG-1999; 99US-0150884P
PR 27-AUG-1999; 99US-0151065P
PR 27-AUG-1999; 99US-0151066P
PR 27-AUG-1999; 99US-0151080P
PR 30-AUG-1999; 99US-0151303P
PR 31-AUG-1999; 99US-0151438P
PR 01-SEP-1999; 99US-0151930P
PR 07-SEP-1999; 99US-0152363P
PR 10-SEP-1999; 99US-0153070P
PR 13-SEP-1999; 99US-0153758P
PR 15-SEP-1999; 99US-0154018P
PR 16-SEP-1999; 99US-0154039P
PR 20-SEP-1999; 99US-0154779P
PR 22-SEP-1999; 99US-0155139P

PR 23-SEP-1999; 99US-0155486P
PR 24-SEP-1999; 99US-0155659P
PR 28-SEP-1999; 99US-0156458P
PR 29-SEP-1999; 99US-0156596P
PR 04-OCT-1999; 99US-0157117P
PR 05-OCT-1999; 99US-0157753P
PR 06-OCT-1999; 99US-0157865P
PR 07-OCT-1999; 99US-0158029P
PR 08-OCT-1999; 99US-0158232P
PR 12-OCT-1999; 99US-0158369P
PR 13-OCT-1999; 99US-0159293P
PR 13-OCT-1999; 99US-0159294P
PR 13-OCT-1999; 99US-0159295P
PR 14-OCT-1999; 99US-0159330P
PR 14-OCT-1999; 99US-0159331P
PR 14-OCT-1999; 99US-0159637P
PR 14-OCT-1999; 99US-0159638P
PR 18-OCT-1999; 99US-0159584P
PR 21-OCT-1999; 99US-0160741P
PR 21-OCT-1999; 99US-0160767P
PR 21-OCT-1999; 99US-0160768P
PR 21-OCT-1999; 99US-0160770P
PR 21-OCT-1999; 99US-0160814P
PR 21-OCT-1999; 99US-0160815P
PR 22-OCT-1999; 99US-0160980P
PR 22-OCT-1999; 99US-0160981P
PR 22-OCT-1999; 99US-0160989P
PR 25-OCT-1999; 99US-0161404P
PR 25-OCT-1999; 99US-0161405P
PR 25-OCT-1999; 99US-0161406P
PR 26-OCT-1999; 99US-0161599P
PR 26-OCT-1999; 99US-0161602P
PR 26-OCT-1999; 99US-0161603P
PR 28-OCT-1999; 99US-0161920P
PR 28-OCT-1999; 99US-0161922P
PR 28-OCT-1999; 99US-0161932P
PR 29-OCT-1999; 99US-0162142P

Query Match 37.0%; Score 1295.5; DB 3; Length 529;
Best Local Similarity 44.6%; Pred. No. 5.7e-111;
Matches 301; Conservative 63; Mismatches 102; Indels 209; Gaps 20;

QY 31 ATSAATVPTFTFNSPQSHLSNFGMCYGVGENGFHSP.TVMP.LKSDGSLCTLEALKSQ 90
DB 22 ATSSSVPPQLVWG--DNTSNFGVCYGSNPGIYSHSVPLRSDGSLCTLEALNRSS 78
QY 91 TCVMVPTSSPKLEDFLGCATMGTHGSHERG--LSLDSIYNSQNAEAPRDLLOSFP 148
DB 79 HSNHODSSPKVEDPFG---THNNNTSHKEMDLSLDSLFYNTH--EPVTTNFOEF 131
QY 149 ---RQGHMSVQTHRYSGLACHGLYQAPLEETFKETHVSDGSLMPOMTGLKNWYA 204
DB 132 FSPFQTRNH-----FEET--RNYGND-----FSLTHG----- 156
QY 205 PTFREFTHOVLQGNQCMGNERNQVSLSGVCCGL-OSLSLSMSPGSSCVAPR-- 261
DB 157 -----GSFNVGV-----YGEFQOSSLNMSPGQSSCLTGSNNH 190
QY 262 -----GTDVAVDAAKRGH---AKGQKQPVHRKSIDT 291
DB 191 QONONQHOSNHQIQISEALVETSVGFETTMAAKKRGQGDVVVGGKQIVHRKSIDT 250
QY 292 FQGRTSQVAGVRRHRTGTGYEAHLMDNSCKEKGCRKRGVOVLTGVDMEEXARAYDLAA 351
DB 251 FQGRTSQVAGVRRHRTGTGYEAHLMDNSFKEGHSRKGRQVLTGVDMEEXARAYDLAA 310
QY 352 LKTMGPSTHINFSIENYQVQLEEMKMSROEYVAHLRRKSGSFSRGASIRGVTRHHQHG 411
DB 311 LKTMGPSTHINFSIENYQVQLEEMKMSROEYVAHLRRKSGSFSRGASIRGVTRHHQHG 370
QY 412 RMOARIGRVAQNKDLYLTGFTQEBAAEAYDYAALKFKGANAVTTFDISRIVERIYMASS 471
```

Db 371 RMOARIGRVAGNKDLYLGTGFTQEEAAEYDVAIKRGTNAVTNPDITRYDVDRIMSN 430  
 Qy 472 NLAGEIARRKNDPRNKIDYKNSVYTSVNNBEETVOQAGNNNNENDEMKCVLFNHP 531  
 Db 431 TLIGELARRR-----NNSIV--VRNTE-----451  
 Qy 532 SQGQOANGNSGDKIMNCNTRNSAFSMALODLIGISVSGQHNMDESSKIGTHPSNT 591  
 Db 452 -----DQALNA-----458  
 Qy 592 SSVTSLSRSREAPSEKRGPSL--LFPMPMETKIV-NPIGTSVTSLPSPVQMRPSPA 648  
 Db 459 ---VVEGSSKNEVSTPEHLSPFPAIFALPQVNOGMFGSNMGNNSPMTSPNAELK-TVA 514  
 Qy 649 ISLSHLPVFASWTD 663  
 Db 515 LTLPMVFPAAMADS 529

RESULT 12  
 ID ABB79640 standard; protein; 585 AA.  
 AC ABB79640;  
 DT 21-OCT-2002 (first entry)  
 DE Cotton AINTEGUMENTA-like polypeptide GHANT1.  
 KW AINTEGUMENTA; ANT-like polypeptide; GHANT1; transgenic plant; cotton;  
 KW plant.  
 OS Gossypium hirsutum.  
 PN WO200259332-A2.  
 PD 01-AUG-2002.  
 PF 19-DEC-2001; 2001WO-US049294.  
 PR 21-DEC-2000; 2000US-0257896P.  
 PA (MONS ) MONSANTO TECHNOLOGY LLC.  
 PI He SS, Dotson SB;  
 DR WPI, 2002-599798/64.  
 PT New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in  
 PT improving agronomic, horticultural, and quality traits of plants, such as  
 PT increased size of plant organs.  
 PS Claim 3; Page 158-161; 169pp; English.

The present sequence is the protein sequence of GHANT1, a newly  
 identified AINTEGUMENTA-like (ANT-like) polypeptide of cotton. The  
 sequence was predicted from a cDNA clone (see ABN84485) isolated in a  
 cotton DNA database screening. The invention provides nucleic acids  
 encoding ANT-like polypeptides comprising, in N-terminal to C-terminal  
 direction, 2 AP2 DNA binding domains followed by an amino acid  
 subsequence selected from those given in ABB79629-35. ANT-like  
 polypeptides were identified in soybean, rice, cotton and corn (see  
 ABB79636-41). Nucleic acids (see ABN84480-86) encoding the ANT-like  
 polypeptides can be used in the construction of transgenic plants,  
 especially corn, soybean, canola, wheat, cotton, tomato or potato (all  
 claimed) having improved agronomic, horticultural or other quality  
 traits, such as increased size of plant organs. These plants are  
 especially useful for production of ethanol or animal feedstuff

Sequence 585 AA;  
 Query Match 33.7%; Score 1180.5; DB 5; Length 585;  
 Best Local Similarity 41.7%; Pred. No. 3.4e-100;

Matches 290; Conservative 79; Mismatches 164; Indels 163; Gaps 25;  
 Qy 17 NWLGSLSPHMKMEATSAATVPTTFYRSPGSHLSFMCMCVGENGFHSPITMPLKS 76  
 Db 3 NWLGSLSPFDLRIDS-----FREDHGFPSS--VMPURS 35  
 Qy 77 DGLSLLEALKRSQT-----QWVPTSSPKLEDELG----GATGTHEY-GS 118  
 Db 36 DGLSLCVDFRRSSIAADEMDYENGIGSATNEGPFLDEPLGCGYSNPSGETAYDGT 95  
 Qy 119 HER-----GLSDSIYNSQNAEAQPN-----RDLISQP--FRQGMNS 155  
 Db 96 HENQNTVPSPRTINNVNAPNYSSGDAEAEENFTNPSFFIQTYNNYENPOTLMAGHSL 155  
 Qy 156 VQTHPYVSLAGHGLYQAPLEETTKETHTVSDSSCLMQMTGKMWVAPTFEFTHOV 215  
 Db 156 QQCDPVPNNQSGVHVHPFESATS-----VSGFSWLRQT-PFPGK-- 197  
 Qy 216 LEOQWNCGMENRNGVSLSGVCGELQSLSLSMSPSQSSCVTAPSGTDSVA---VDK 271  
 Db 198 -----ASGNETNNNF-----NFOALSLTMSF-----TSRNGFPALAPLEVDR 236  
 Qy 272 KSGHATLGQKQVYRKRKSIDTEGQRTSOYRGVTRHRMTGYENHMDNCKGGRKGRQ 331  
 Db 237 KRPVGNLTFRESVPRKSIDTFQORTSOYRGVTRHRMTGYEHLMDNCKRGGRKGRQ 296  
 Qy 332 VYLGVDMEKAKARAYDLAALKYMGSPSTHINFSIENYCYOLEMKMSROEYVAHLRRKS 391  
 Db 297 VYLGVDKEKAKAYDLAALKYMGSTHINPLSTYKELEEMKQMRQEVVALRRKS 356  
 Qy 392 SGFSRGASLYRGVTRHQRHQRWQARIGRYAGNKDLYLGFSTQEEAAEYDVAIKFRGA 451  
 Db 357 SGFSRGASLYRGVTRHQRHQRWQARIGRYAGNKDLYLGFSTQEEAAEYDVAIKFRGT 416  
 Qy 452 NATNFDIARYDYERIMASSNLLAGELARRKNDPRANDIDYKNSVTSVNNBEETVOQ 511  
 Db 417 SALTNPDIRYDYKRCSSSTLIGELARRKNDPRANDIDYKNSVTSVNNBEETVOQ 475  
 Qy 512 AGNNNNE-NDSEWKVYLFNHPGQQ-QANGSGDQKIMNCYKRNASFMALQDLIGT-- 567  
 Db 476 SGASDELADMTWTA---NSDEQCHOSTNTNNDASLANSS-RUSSNPQSPKSGGLAS 531  
 Qy 568 DSVSGQHNMDESSKIGTHPSNTSLVTSLSRSREAPSEKRGPSLFPMPMETKIVNP 627  
 Db 532 DKFGIG---GDYSHHG-YFSLKSGKYEDGNSSTNSVNR-----567  
 Qy 628 IGTSVTSLPSPVQMRPSPAISLSH-LPFAWTD 662  
 Db 568 LGN-----LGLVHKIMPFALMN2 585

RESULT 13  
 ID ABR40856 standard; protein; 557 AA.  
 AC ABR40856;  
 DT 16-MAY-2003 (first entry)  
 DE Glycine max oil trait related protein sequence SEQ ID NO:491.  
 KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;  
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;  
 KW lipid-like transcription factor caleosin; ATP citrate lyase; SNF1;  
 KW CMC-like transcription factor; antisense inhibition; co-suppression;  
 KW transgenic plant.  
 OS Glycine max.  
 PN WO2003002751-A2.  
 PD 09-JAN-2003.



PF 27-JUN-2002; 2002W0-US020152.  
 XX  
 PR 29-JUN-2001; 2001US-0301913P.  
 XX  
 XX (DUPLO) DU PONT DE NEMOURS & CO E. I.  
 PA (PION-) PIONEER HI-BRED INT INC.  
 PA  
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harwell LT,  
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,  
 PI Tarczyński MC;  
 XX  
 DR WPI: 2003-201509/19.  
 DR N-PSDB; ACC00850.  
 XX  
 PT Novel nucleotide fragment encoding polypeptides having receptor-like  
 PT protein kinase activity, caleosin-like activity, useful for altering oil  
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.  
 XX  
 PS Claim 12; Page 509-511; 542pp; English.  
 XX  
 CC The present invention describes an isolated nucleotide fragment (I)  
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a  
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen  
 CC activated protein (MAP)-kinase activity, lipid-like transcription factor  
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like  
 CC activity and CXC-like transcription factor activity. Also described: (1)  
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or  
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)  
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)  
 CC oil obtained from (V). (I) or its part can be used in antisense  
 CC inhibition or co-suppression in a transformed plant. (III) is useful for  
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,  
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for  
 CC creating transgenic plants having altered lipid profiles. (I) can also be  
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to  
 CC ABR40879 represent sequences used in the exemplification of the present  
 CC invention  
 CC  
 SQ Sequence 557 AA;

Query Match 30.3%; Score 1061.5; DB 6; Length 557;  
 Best Local Similarity 43.1%; Pred. No. 3.8e-89;

Matches 260; Conservative 67; Mismatches 125; Indels 151; Gaps 23;

QY 100 PLEDEFLGATGATGTHYSGHERGLSDLSIYNSSONA-----BAQPNRDLLOQPFROQGM 154  
 |||||  
 DB 66 PKLEDFLG-----CYSDSPAKYFCQDSQPDQ--NQSONNYSKI 101  
 QY 155 SVQTHYVYSGLA CHGLYQAPLEETTKETHVSDCSLMPQW-----TEGL-KN 201  
 |||||  
 DB 102 NNVAAPSF-----CTN-----DSELETGNTLPSLSLSPHAYNDNHALPTNGMKXS 151  
 QY 202 WVAPEFSTHQVLEEQMNCGMERNNGVSLGVCGLQSLSLNSPGSSQSCVTAPS 261  
 |||||  
 DB 152 WLAQT-QFSDDK-----PSNEANGCN-----FQSLSTLMSPSVQNG-VGALS 192  
 QY 262 GIDSDVAVDAKKGHAGLGGOKOPYHRKSIDTFEGQRTSOVRGVTGRHWTGRYAHLMDSCK 321  
 |||||  
 DB 193 SV-QVNEDESKRYMAKSHAPEVPRKSIDTFQRTSOVRGVTGRHWTGRYAHLMDSCKR 251  
 QY 322 KEGQTKRGQVYLGADYMEKAKARAYDALALKYMGSPSTHINSINYQVLEEMKMSQ 381  
 |||||  
 DB 252 KEGQTKRGQVYLGADYMEKAKARAYDALALKYMGSPSTHINSINYQVLEEMKMSQ 311  
 QY 382 EYVAHTRKSSGFRASIRYGVTRHNGRQWQARIGRAYNKDYLIGFTSGQEAARAY 441  
 |||||  
 DB 312 EFVANNRRKSSGFRASIRYGVTRHNGRQWQARIGRAYNKDYLIGFTSGQEAARAY 371  
 QY 442 DVAAIKFRGANAVTNDISRYDVERIMASSNLLAGELARRKDNPPNKDI-DYN---KS 497  
 |||||  
 DB 372 DVAIKFRGTSVANTNDISRYDVKRICSSSTLLAGELARRSKESPAPVADDFMSQSS 431  
 QY 498 VVTSVNNETVQVQAGNNNNE-NDSEWRKVLFNHPSQQQAGANGSSDQKINAGCYRNSA 556

DB 432 PWEPLVSQPPPLAITDGEHSDLSNNMM-----NANNSDDAQN----- 469  
 QY 557 FSWALQDLIGIDVSGGQHNMLDESSKIGTHFSNTSLVTS:SSSREASPERKGPGLFP 616  
 DB 470 -----ESG--GAEFNNN---VTSSSSSQOVSPSS-NKDALNP 500  
 QY 617 MPPEMTKIVNPIGTSVTSW-----LPSPVQ-----KSPSPAIS:SHLPVPFAS 659  
 DB 501 QSP-----NEFGVSADYGHGYFTLDGPKYDDGNNNDMSNTRNLGNLGVNPMPRL 554  
 QY 660 WTD 662  
 DB 555 MNE 557  
 RESULT 14  
 ABR40768  
 ID ABR40768 standard; protein; 710 AA.  
 XX  
 AC ABR40768;  
 XX  
 DT 16-MAY-2003 (first entry)  
 DE Glycine max oil trait related protein sequence SEQ ID NO:356.  
 XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;  
 XX receptor-like protein kinase; mitogen activated protein kinase; oil;  
 XX lipid-like transcription factor caleosin; ATP citrate lyase; SNF1;  
 XX CXC-like transcription factor; antisense inhibition; co-suppression;  
 XX transgenic plant.  
 XX  
 OS Glycine max.  
 XX  
 XX  
 PN W02003002751-A2.  
 XX  
 PD 09-JAN-2003.  
 XX  
 XX 27-JUN-2002; 2002W0-US020152.  
 PF  
 PR 29-JUN-2001; 2001US-0301913P.  
 XX  
 XX (DUPLO) DU PONT DE NEMOURS & CO E. I.  
 PA (PION-) PIONEER HI-BRED INT INC.  
 PA  
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harwell LT,  
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,  
 PI Tarczyński MC;  
 XX  
 DR WPI: 2003-201509/19.  
 DR N-PSDB; ACC00803.  
 XX  
 PT Novel nucleotide fragment encoding polypeptides having receptor-like  
 PT protein kinase activity, caleosin-like activity, useful for altering oil  
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.  
 XX  
 PS Claim 12; Page 374-376; 542pp; English.  
 XX  
 CC The present invention describes an isolated nucleotide fragment (I)  
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a  
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen  
 CC activated protein (MAP)-kinase activity, lipid-like transcription factor  
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like  
 CC activity and CXC-like transcription factor activity. Also described: (1)  
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or  
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)  
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)  
 CC oil obtained from (V). (I) or its part can be used in antisense  
 CC inhibition or co-suppression in a transformed plant. (III) is useful for  
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,  
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for  
 CC creating transgenic plants having altered lipid profiles. (I) can also be  
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to

CC ABR40879 represent sequences used in the exemplification of the present  
 CC invention  
 XX  
 XX Sequence 710 AA;

Query Match 30.3%; Score 1059.5; DB 6; Length 710;  
 Best Local Similarity 36.8%; Pred. No. 8,8e-89;  
 Matches 284; Conservative 74; Mismatches 203; Indels 211; Gaps 24;

QY 17 NMLGSLSPMKKAEAT--AAVPTTFMSP-----GSHLSNFGMXY 57  
 DB 5 NLGFLSLPHEHPSSQDHSQTPSRFCFNPDSISTDVAGCPDLSDSTPHLNL-PSY 63  
 QY 58 GVGE-----NGNFHSPITVMPKSD-----GSLCILEALKRSQTVWVPTSSPLEDF 105  
 DB 64 GIVAFHRSNNIH---TTQDWKENYNSQNLGTSQNNQNNANHQOQO---CQCPYLENF 118  
 QY 106 LGGATWGTHERYSGHERGLSLDSIYNSQNAEAPNDLSPFRQCGHMSVQTHPYYSGL 165  
 DB 119 LGG-----HSRGEHE-----Q-----YQGN 134  
 QY 166 ACHGLYQAPLEBEET-----TKETHVDCSSIMPQTEGLKXWVAPTEFSTHQVLRQ 218  
 DB 135 SASRTYMPAPVPLAGGGGGGSSNSNTSSISGLSM---IKTWL---RNPDSHNNNN 188  
 QY 219 QMNGCMGNERNGVSLGSGCGELQSLSLMSPGSSQSCVTAPSGTDSVAVDAKKRGHAKL 278  
 DB 189 NNNSSGGNSRSVSQ-----QTLISMSSTGQSS--TSLPILTASV---DNSESS 233  
 QY 279 GQKQP-----VHKSIDTFGQRTSOVGVTRRMTGRYFAHLMDSCKK 322  
 DB 234 DNKQPHHTALDITQTGALIEPAKPSIDTFGQRTSITRGVTRHMTGRYFAHLMDSCKR 293  
 QY 323 EGQTRKGRQVYLGQYDMEKKAARAADALAKYWPSTHINPSLENYVQLEKNNKSRQ 382  
 DB 294 EGQTRKGRQVYLGQYDMEKKAARAADALAKYWPSTHINPSLENYVQLEKNNKSRQ 353  
 QY 383 YVAHLRRKSSGFSGASIRGVTRHHQHWQARIGVAGNKDLYLGTSTQEEAAEAYD 442  
 DB 354 YVASLRKSSGFSGASIRGVTRHHQHWQARIGVAGNKDLYLGTSTQEEAAEAYD 413  
 QY 443 VAAIKFGANAVTFDISRYDVERIMASSNLLAGELARRKKD-----NDPKNCDIDY 494  
 DB 414 VAAIKFRLSAVTFDMGRYVKSILBETPLIGAKAKELKDMQVEYLRVENVHRAQED 473  
 QY 495 NKSVVTS-----VNNEETVOVQANNNNENDSEWKKVLFNHPQOQOQANNGSDQKIM 548  
 DB 474 HSSIMNSHLLQGIINN---YAAAGTTATHHHNNHNAALAFHQPCTTIHYPGQRINW 528  
 QY 549 CGNTR---NSAFPMALODLIGIDSVSGGCHMLDESSKIGTHFNTSSLYTSLSSREAS 605  
 DB 529 CKQEDNSDASHSLYSIDHQLQGNNTNHPFTNSGLHMLMDASIDNSSSSSVV 588  
 QY 606 PEKRGPSLLFPMPMETKIV-----NP----- 627  
 DB 589 YDVGYGGGGVYVIPMGITTTTVVANDGQNPSSNHGFDNEIKALGYESVGSITDPYHAK 648  
 QY 628 -----IGTSVTMLPS--PTVQMRPSPAISLH 653  
 DB 649 ARLNYLLTQOQSSVDAYKASAYDQSACTWVPAITTAHPRSTSMALCH 700

RESULT 15  
 ABR40766  
 ID ABR40766 standard; protein; 707 AA.  
 XX  
 XX ABR40766;  
 AC  
 XX  
 XX 16-MAY-2003 (first entry)  
 DT  
 XX  
 XX Glycine max oil trait related protein sequence SEQ ID NO:352.  
 DE  
 XX  
 XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;

KW receptor-like protein kinase; mitogen activated protein kinase; oil;  
 KW lipid-like transcription factor caleosin; ATP citrate lyase; SNF1;  
 KW CKC-like transcription factor; antisense inhibition; co-suppression;  
 KW transgenic plant.

OS Glycine max.

PN WO2003002751-A2.

PD 09-JAN-2003.

PF 27-JUN-2002; 2002WO-US020152.

PR 29-JUN-2001; 2001US-0301913P.

PA (DUPO) DU PONT DE NEMOURS & CO B. I.

PI (PION-) PIONEER HI-BRED INT INC.

PI Allen SM, Allen MB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT,  
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,  
 PI Tarczynski MC;

PI WPI; 2003-201509/19.

DR N-PSDB; ACC00801.

PT Novel nucleotide fragment encoding polypeptides having receptor-like  
 PT protein kinase activity; caleosin-like activity; useful for altering oil  
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.

PS Claim 12; Page 369-371; 542pp; English.

XX The present invention describes an isolated nucleotide fragment (I)  
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a  
 CC polypeptide (PP) having receptor-like protein kinase activity; mitogen  
 CC activated protein (MAP)-kinase activity; lipid-like transcription factor  
 CC activity; caleosin-like activity; ATP citrate lyase activity; SNF1-like  
 CC activity; and CKC-like transcription factor activity. Also described: (1)  
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or  
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)  
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)  
 CC oil obtained from (V). (I) or its part can be used in antisense  
 CC inhibition or co-suppression in a transformed plant. (III) is useful for  
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,  
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for  
 CC creating transgenic plants having altered lipid profiles. (I) can also be  
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to  
 CC ABR40879 represent sequences used in the exemplification of the present  
 CC invention

SQ Sequence 707 AA;

Query Match 29.9%; Score 1046.5; DB 6; Length 707;

Best Local Similarity 36.8%; Pred. No. 1.4e-87;

Matches 285; Conservative 88; Mismatches 182; Indels 219; Gaps 30;

QY 17 NMLGSLSPMKKAEAT--AAVPTTFMSP-----GSHLSNFGMXY 55  
 DB 5 NLGFLSLPHEHPSSQDHSQTPSRFCFNPDSISTDVAGCPDLSDSTPHLNL-P 63  
 QY 56 CYGVGNGNPH---SPLTVAPLKSD-----GSLCILEALKRSQTVWVPTSSPLEDF 104  
 DB 64 SYGIYE--AFHNNSNTITQDWKENYNSQNLGTSQNNQNNANHQOQO---QPKLEN 115  
 QY 105 FLGGATWGTHERYSGHERGLSLDSIYNSQNAEAPNDLSPFRQCGHMSVQTHPYYSGL 164  
 DB 116 FLGG-----HSRGEHE-----Q-----YQGN 131  
 QY 165 LACHGLYQAPLEBEET-----KETHVDCSSIMPQTEGLKXWVAPTEFSTHQVLRQ 217  
 DB 132 NSASTYMPAPVPLAGGGGGGSSNSNTSSISGLSM---IKTWL---RNPDSHNNNN 181  
 QY 218 QMNGCMGNERNGVSLGSGCGELQSLSLMSPGSSQSCVTAPSGTDSVAVDAKKRGHAK 277

Wed Mar 17 08:18:54 2004

us-10-024-632-2.rag

Page 16

Db 182 ENIN-----NNESGNIRSSVQOTLSLSMSTGSSS-TSLPLLTASY-----DNCESS 229  
Qy 278 LGOKCP-----VHRKSIDTFCGRTSOYRGVTRHRMTGRYEALHWDNSCK 321  
Db 230 SDKQOPNTSAALDSTOGALETAPRKSIDTFCGRTSIRGVTRHRMTGRYEALHWDNSCR 289  
Qy 322 KEGOTRKGROYTIGGYMEKABAYDLALKTWGSPSTHINSIENYOQLEBKMSRQ 381  
Db 290 REGOTRKGROYTIGGYMEKABAYDLALKTWGSPSTHINSIENYOQLEBKMSRQ 349  
Qy 382 EYVAHLRKRKSGFSGASIRGVTRHHRHGRMOCARIGRVAGNKDLYLGFTSTOEBAAEAY 441  
Db 350 EYVASLARKSGFSGASIRGVTRHHRHGRMOCARIGRVAGNKDLYLGFTSTOEBAAEAY 409  
Qy 442 DYAAIKFRGANVTNPFISRIVDERIMASSNLAGELARKK-----DNDRPKDIT 492  
Db 410 DYAAIKFRGLSAVTNPFMSRYDVKSILESTTLPIGGAARLKDMEQVELSVDNHGRADQV 469  
Qy 493 DYNKSVVTS-----VNNETVOVQAQNNNNNENDSEMKM-----VLFNHP-----SQO 534  
Db 470 DH--SITMSHLPQGINNNYAGGTAETHNWHNAHAFHQPOCTMHPYGGRIWCKQE 527  
Qy 535 QOANGNG-----SDOKIMNCNYRNSAF---SMALQDIDISV-----570  
Db 528 QOONSZAPHSLSYSDIHQLOLGNNGTNEFHTNGLHPWLSMDSASIDNSSSSNSVYDG 587  
Qy 571 --GSGQHML-----DESSKIGTHFSNT-----SSLVTSLSSSRFASPEK 608  
Db 588 YGGGGGYNNVMPWGTTAIVVASDGDQNPNSNHGFGNEIKALGYESVYGSATDSYAHARN 647  
Qy 609 RGSBLFPMPMETKIYVPI-----GTSVTSMLPS--PTVQMRPSPALISLH 653  
Db 648 ----LYYLTQOQSSSVYDVKASAYDQGSACNWTVPALPTAHPRSTYSMALCH 696

Search completed: March 9, 2004, 10:45:34  
Job time : 63 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using ew model

Run on: March 9, 2004, 10:44:24; Search time 23 Seconds

(without alignments) 1488.175 Million cell updates/sec

Title: US-10-024-632-2

Perfect score: 3502

Sequence: 1 MKRINESNNTDDGNHNLWG.....RSPALISHLFVFASTWTD 663

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued Patents AA.\*

1: /cgnt2\_6/prodata/2/1aa/5A\_COMB.pep.\*  
2: /cgnt2\_6/prodata/2/1aa/5B\_COMB.pep.\*  
3: /cgnt2\_6/prodata/2/1aa/6A\_COMB.pep.\*  
4: /cgnt2\_6/prodata/2/1aa/6B\_COMB.pep.\*  
5: /cgnt2\_6/prodata/2/1aa/PCTUS\_COMB.pep.\*  
6: /cgnt2\_6/prodata/2/1aa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1328	37.9	555	US-09-227-421-2	Sequence 2, Appl
2	1328	37.9	555	US-09-479-855-2	Sequence 2, Appl
3	1307	37.3	548	US-09-227-421-5	Sequence 5, Appl
4	1307	37.3	548	US-09-479-855-5	Sequence 5, Appl
5	517	14.8	432	US-08-700-152A-4	Sequence 4, Appl
6	385	11.0	77	US-08-912-272-9	Sequence 9, Appl
7	385	11.0	77	US-09-026-039-9	Sequence 9, Appl
8	346	9.9	69	US-08-912-272-10	Sequence 10, Appl
9	346	9.9	69	US-09-026-039-10	Sequence 10, Appl
10	197	5.6	67	US-08-700-152A-1	Sequence 1, Appl
11	192.5	5.5	67	US-08-912-272-11	Sequence 11, Appl
12	192.5	5.5	67	US-09-026-039-11	Sequence 11, Appl
13	191.5	5.5	68	US-08-700-152A-2	Sequence 2, Appl
14	191.5	5.5	68	US-08-912-272-5	Sequence 5, Appl
15	191.5	5.5	68	US-09-026-039-5	Sequence 5, Appl
16	188	5.4	67	US-08-912-272-4	Sequence 4, Appl
17	188	5.4	67	US-09-026-039-4	Sequence 4, Appl
18	156.5	4.5	248	US-09-202-161B-3	Sequence 3, Appl
19	153	4.4	251	US-09-198-119C-55	Sequence 55, Appl
20	148.5	4.2	251	US-09-198-119C-87	Sequence 87, Appl
21	147	4.2	248	US-09-533-029-22	Sequence 22, Appl
22	146	4.2	250	US-09-198-119C-71	Sequence 71, Appl
23	145	4.1	250	US-09-198-119C-53	Sequence 53, Appl
24	141.5	4.0	344	US-09-533-029-66	Sequence 66, Appl
25	140.5	4.0	371	US-09-533-029-82	Sequence 82, Appl
26	138.5	4.0	375	US-09-533-029-82	Sequence 82, Appl
27	137.5	3.9	328	US-09-300-672-12	Sequence 12, Appl

28	136	3.9	210	US-09-198-119C-17	Sequence 17, Appl
29	136	3.9	210	US-09-198-119C-47	Sequence 47, Appl
30	135.5	3.9	255	US-09-198-119C-83	Sequence 83, Appl
31	135	3.9	231	US-09-202-161B-1	Sequence 1, Appl
32	134.5	3.8	58	US-09-202-161B-27	Sequence 27, Appl
33	134.5	3.8	279	US-09-198-119C-51	Sequence 51, Appl
34	134	3.8	243	US-09-533-029-34	Sequence 34, Appl
35	133.5	3.8	280	US-09-198-119C-77	Sequence 77, Appl
36	133	3.8	209	US-09-198-119C-93	Sequence 93, Appl
37	132.5	3.8	66	US-08-912-272-29	Sequence 29, Appl
38	132.5	3.8	66	US-09-026-039-29	Sequence 29, Appl
39	132.5	3.8	277	US-09-198-119C-61	Sequence 61, Appl
40	132.5	3.8	277	US-09-198-119C-69	Sequence 69, Appl
41	132.5	3.8	283	US-09-198-119C-63	Sequence 63, Appl
42	132	3.8	252	US-09-198-119C-91	Sequence 91, Appl
43	131.5	3.8	264	US-08-894-731-4	Sequence 4, Appl
44	131	3.7	213	US-09-198-119C-59	Sequence 59, Appl
45	130.5	3.7	277	US-09-198-119C-57	Sequence 57, Appl

## ALIGNMENTS

RESULT 1					
US-09-227-421-2					
Sequence 2, Application US/09227421					
Patent No. 6559357					
GENERAL INFORMATION:					
APPLICANT: Fischer, Robert L.					
APPLICANT: Mizukami, Yukiko					
TITLE OF INVENTION: The Regents of the University of California					
TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility					
FILE REFERENCE: 023070-090700PC					
CURRENT APPLICATION NUMBER: US/09/227,421					
PRIOR FILING DATE: 1999-01-08					
PRIORITY APPLICATION NUMBER: US 09/227,421					
NUMBER OF SEQ ID NOS: 8					
SOFTWARE: Patent In Ver. 2.1					
SEQ ID NO 2					
LENGTH: 555					
TYPE: PRT					
ORGANISM: Arabidopsis thaliana					
US-09-227-421-2					
Query Match					
Best Local Similarity 37.9%, Score 1328, DB 4, Length 555,					
Best Local Similarity 43.5%, Pred. No. 3,4e-123,					
Matches 314, Conservative 69, Mismatches 113, Indels 226, Gaps 22;					
QY	1	MKRINESNNTDDGNHNLWG.....EATSAVPTTFYM	43		
DB	1	MKSPDNDNNHSTNLTILGSLSSNMKGGKGRALYSSTSSAATSSSVPLV	60		
QY	44	SPSGHLSNFMCYGVGENKTFSPPLVMPKSDGICLIALKRSGQVMPVTSFKL	103		
DB	61	G---DNTSNFVGVCGSNPNNGIYSHMSVMPLRSGSLCLNEALNRSSNKHODSSPKYE	117		
QY	104	DFLGATGTGTREYSHRG--LSLDSIYNSQNAEAPNDLSQPF---ROOGHVSQ	157		
DB	116	DFPG---THNNNTHHEANDLSLDSLFYVTTT---EPNTTNPQEFSPQVRN----	166		
QY	158	THPIYSGIACGLYQAPLEETTTKETHVSCSSLMPTGGLKNWVAPTREFTHQVLE	217		
DB	167	-----EET--RNYGND-----PGLTGG-----	182		
QY	218	QMNCGMNEBNGVSLGSGGEL--QSLSSMSPGSSQSCVTAPS-----	261		
DB	193	-----GTFVGV-----YGFQGLSLSSMSPGSSQSCITSHHQNQNONHOSQKH	229		
QY	262	-----GTSVAVDKAKRGH---AKLGQKQPVHRKSITLFGQRTSQYFGVTR	304		
DB	230	QOISEALVETSVGFEITTTAAAKKRGQEDVVVVQKQIVHRKSIDTFQGRISQYRGVTR	289		

305 HMTGTYEAHLMDNSCKKEGQTRKGRQVYLLGQYDMEERKAAAYDLAALKTWGPSTHINF 364  
290 HMTGTYEAHLMDNSCKKEGQTRKGRQVYLLGQYDMEERKAAAYDLAALKTWGPSTHINF 349  
365 IENVOUAEEMKNSROEYVAHLRRKSGFSRGSITRGVTRHGHGRQARIGRVAANK 424  
350 AENYQXELIDMKMTRQETVAHLRRKSGFSRGSITRGVTRHGHGRQARIGRVAANK 409  
425 DLYLGTFTGEBAAEAYDVAALIKFRGANAVTNFDISRYDVERIMASSNLLAGELARRKD 484  
410 DLYLGTFTGEBAAEAYDVAALIKFRGTNAVTNFDITRYDVRIMSSNTLLSGELARR-- 467  
485 NDPRKCIDLYNKSIVTYSNNETTYVQAGNNNNENDESEKAVLFPHNSQQQOQANGSDQ 544  
468 -----NNSIV--VRNTE-----DQ 479  
545 KIMNGCYRNBSAFSVALODLIGIDSVGSGQHMLDESSKIGTHFNTSSLYTSSSREA 604  
480 TALNA-----VEGGSNKEY 494  
605 SPEKRGPSL--LFPYPMETKIV-NPIGTSVTSMLEPPTVQMRPSPALISHLPPFASWT 661  
495 STEPLRISFPALFALPQVQKMFSGNMGNNMSPMTSNPAELK-TVALTLPPMPVFAAMA 553  
662 DT 663  
554 DS 555

RESULT 2  
US-09-479-855-2  
Sequence 2, Application US/09479855  
Patent No. 6639128  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Mizukami, Yukiko  
TITLE OF INVENTION: The Regents of the University of California  
TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility  
FILE REFERENCE: 023070-090720US  
CURRENT APPLICATION NUMBER: US/09/479,855  
CURRENT FILING DATE: 2000-01-07  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 555  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
OTHER INFORMATION: AINTEGUMENTA (ANT)  
US-09-479-855-2

Query Match 37.9%; Score 1328; DB 4; Length 555;  
Best Local Similarity 43.5%; Pred. No. 3,4e-123;  
Matches 314; Conservative 69; Mismatches 113; Indels 226; Gaps 22;

1 MKRIVESNNITDDGNNHMLGFSISPH-MKM-----EATSAATYPTTFM 43  
1 MKSPCDNDNNHNTTNLLGFSISSNMKMGGRGREALYSSSTSAATSSSVPLVV 60  
44 SPQSGHLSFGMPCYGVGENGFHSPLTYWPLKSDGSLCILEALRSQTVWVPSPKLE 103  
61 G---DNTSNFGVCGSNNGIYSHMSYMPLRSDGSLCLMALNRSSHHHODSSPKVE 117  
104 DFLGATWGTHERYSGHERG--LSIDSIYNSQNAEAPNRDLISQF-----RQGGHNSVQ 157  
118 DFFG---THNNNTSHKAMDLSDSLFYNTTH--EENTTNTQEEFFSPQCTNH--- 166  
158 THPYSGLAGHLYQAPLESEETKETHYSDCSLMPQWTEGLKNMVAPTREFTTHQVLE 217  
167 -----EEET--RNYGND-----PSLTHG----- 182  
218 QQANCGMGNERNGVLSVGGCEL-QSLISLMSFSGOSSCVTAPS----- 261

183 -----GSFNVGV-----YGEFOQSLISMSFSGOSSCITGSHHNOONONHOSQNH 229  
262 -----GTDVAVADAKRGH---AKLGQXQPVHRKSIDTFGORTSOYRGVTR 304  
230 QQISEALVETSVGFETTTMAAKKRGQEDVYVQKQIVHRKSIDTFGORTSOYRGVTR 289  
305 HMTGTYEAHLMDNSCKKEGQTRKGRQVYLLGQYDMEERKAAAYDLAALKTWGPSTHINF 364  
290 HMTGTYEAHLMDNSCKKEGQTRKGRQVYLLGQYDMEERKAAAYDLAALKTWGPSTHINF 349  
365 IENVOUAEEMKNSROEYVAHLRRKSGFSRGSITRGVTRHGHGRQARIGRVAANK 424  
350 AENYQXELIDMKMTRQETVAHLRRKSGFSRGSITRGVTRHGHGRQARIGRVAANK 409  
425 DLYLGTFTGEBAAEAYDVAALIKFRGANAVTNFDISRYDVERIMASSNLLAGELARRKD 484  
410 DLYLGTFTGEBAAEAYDVAALIKFRGTNAVTNFDITRYDVRIMSSNTLLSGELARR-- 467  
485 NDPRKCIDLYNKSIVTYSNNETTYVQAGNNNNENDESEKAVLFPHNSQQQOQANGSDQ 544  
468 -----NNSIV--VRNTE-----DQ 479  
545 KIMNGCYRNBSAFSVALODLIGIDSVGSGQHMLDESSKIGTHFNTSSLYTSSSREA 604  
480 TALNA-----VEGGSNKEY 494  
605 SPEKRGPSL--LFPYPMETKIV-NPIGTSVTSMLEPPTVQMRPSPALISHLPPFASWT 661  
495 STEPLRISFPALFALPQVQKMFSGNMGNNMSPMTSNPAELK-TVALTLPPMPVFAAMA 553  
662 DT 663  
554 DS 555

RESULT 3  
US-09-227-421-5  
Sequence 5, Application US/09227421  
Patent No. 6559357  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Mizukami, Yukiko  
TITLE OF INVENTION: The Regents of the University of California  
TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility  
FILE REFERENCE: 023070-090700PC  
CURRENT APPLICATION NUMBER: US/09/227,421  
CURRENT FILING DATE: 1999-01-08  
PRIOR APPLICATION NUMBER: US 09/227,421  
PRIOR FILING DATE: 1999-01-08  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 5  
LENGTH: 548  
TYPE: PRT  
ORGANISM: Brassica napus  
US-09-227-421-5

Query Match 37.3%; Score 1307; DB 4; Length 548;  
Best Local Similarity 43.6%; Pred. No. 4.1e-121;  
Matches 309; Conservative 77; Mismatches 103; Indels 220; Gaps 25;

9 NITDDGNNHMLGFSISPHM-----KVENSAATVPTTFMSPBQ-----SHLSFGMXY 57  
6 DNDSDVTTNLLGFSISSNMKMGGRGREALYSSSSSVATSS--SVPLVVGDNSSMYGVCY 63  
58 GVG-ENGNFHSPLYTWPPLKSDGSLCILEALKRS-----QTVWVPSPLEDFLGA 109  
64 GSNLAARENYQMSYMPLRSDGSLCLMALNRSSHHHNSOV-----SSPMEBF----- 115  
110 TWGTHRYG-SHERG--LSIDSIYNSQNAEAPNRDLISQFPPROGSHSVQTHPYSGLA 166

Db 116 -FGTHNHTSHKEMDLSLDSLFYNTTHA---PNNNTNFOEF---FSPQTRNNH--- 163  
 Qy 167 CHGLYQALBEEETTKETHVSDCSLMPQMTGKNNVAFPRFSTHQVLEQCMGCMGN 226  
 Db 164 -----BEET--RNYEND-----PGLTHG----- 179  
 Qy 227 ERNGVSLGVSOGCEL--QSLSMSPGSQSSCVTA-----P 260  
 Db 180 ---GGSFNVGVYGEFQOSLSLSMSPGSSCITASHHQNQTONHOQISEALVETSAFGE 236  
 Qy 261 SGTDSVAVDACKRG--HAKLGOKOPYHRSIDTFQORTSOYRGVTRHRTMGRVYAHLMDS 319  
 Db 237 TTTMAAAAKKRGQEVVGVGQVYHRSIDTFQORTSOYRGVTRHRTMGRVYAHLMDS 296  
 Qy 320 CKKEGQTRKGRQVYLGVDMEKARAYDLAALKYWGSPSTHINSIENYQVLEBKMS 379  
 Db 297 FKKEGHSKRGQVYLGVDMEKARAYDLAALKYWGSPSTHINSIENYQVLEBKMS 356  
 Qy 380 ROEYVAHLRRKSSGFSRGSASTYRGVTRHGHQWQARIGVAGNKDLYLGTFQOEBAAE 439  
 Db 357 ROEYVAHLRRKTSGFSRGSASTYRGVTRHGHQWQARIGVAGNKDLYLGTFQOEBAAE 416  
 Qy 440 AYDVAALFKRGANAVTNFDISRVDYERIMASSNLLAGELARRKKNDPRKIDYKSAVY 499  
 Db 417 AYDVAALFKRGANAVTNFDISRVDYERIMASSNLLAGELARRKKNDPRKIDYKSAVY 465  
 Qy 500 TSVNNEE---TVQVQAGNNNNENDESMKVLFNHPSQQQOANGSDQKIMNCGNTRNSA 556  
 Db 466 RNISDEEALTAIVVAGSNKEV-----GSPERVLSF----- 496  
 Qy 557 FSMALQDLIGDVSQGHNMLEDSSKIGTHFSNTSLVTSLSRSRASPBRKGPILFP 616  
 Db 497 -----PTIFA 501  
 Qy 617 MPMEETKI--VNPIGTSVSWLPSPPTVQMRPSPALSHLPVFAWTD 663  
 Db 502 LPQVGPXMGANVVG--NMSWTTNPNADLK--TVSLTLQPMVFAWADS 548  
 RESULT 4  
 US-09-479-855-5  
 / Sequence 5, Application US/09479855  
 / Patent No. 6639128  
 / GENERAL INFORMATION:  
 / APPLICANT: Fischer, Robert L.  
 / APPLICANT: Mizukami, Yukiko  
 / APPLICANT: The Regents of the University of California  
 / TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility  
 / TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants  
 / FILE REFERENCE: 023070-090720US  
 / CURRENT APPLICATION NUMBER: US/09/479,855  
 / CURRENT FILING DATE: 2000-01-07  
 / NUMBER OF SEQ ID NOS: 8  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 5  
 / LENGTH: 548  
 / TYPE: PRT  
 / ORGANISM: Brassica napus  
 / OTHER INFORMATION: canola AINTEGUMENTA (ANT)  
 US-09-479-855-5

Query Match 37.3%; Score 1307; DB 4; Length 548;  
 Best Local Similarity 43.6%; Pred. No. 4,1e-121;  
 Matches 309; Conservative 77; Mismatches 103; Indels 220; Gaps 25;

Qy 9 NTDDGANNWLGFSLSPH-----KMEATSAATVPTFYMSPO-----SHLSNFGMCY 57  
 Db 6 DNDDNTNLLGFSLSNMLKMGGEALYSSSSSVATS--SVFPQVWGNSNNTGVCY 63  
 Qy 58 GVG-ENGNFHSPPLVTPMKISDGLICILEALRS-----QTVWVPTSSPKLEDFLGA 109  
 Db 64 GSNLAAREMYGQMSVMPLRSDGLICIMBALNRSSSHSNHHHSQV-----SSPKMEDF----- 115

Qy 110 TWGTHEYG-SHERG--ISLDSIYNSQNAEQPNRDLISQFROQGHMSVQTHPYSGLA 166  
 Db 116 -FGTHNHTSHKEMDLSLDSLFYNTTHA---PNNNTNFOEF---FSPQTRNNH--- 163  
 Qy 167 CHGLYQALBEEETTKETHVSDCSLMPQMTGKNNVAFPRFSTHQVLEQCMGCMGN 226  
 Db 164 -----BEET--RNYEND-----PGLTHG----- 179  
 Qy 227 ERNGVSLGVSOGCEL--QSLSMSPGSQSSCVTA-----P 260  
 Db 180 ---GGSFNVGVYGEFQOSLSLSMSPGSSCITASHHQNQTONHOQISEALVETSAFGE 236  
 Qy 261 SGTDSVAVDACKRG--HAKLGOKOPYHRSIDTFQORTSOYRGVTRHRTMGRVYAHLMDS 319  
 Db 237 TTTMAAAAKKRGQEVVGVGQVYHRSIDTFQORTSOYRGVTRHRTMGRVYAHLMDS 296  
 Qy 320 CKKEGQTRKGRQVYLGVDMEKARAYDLAALKYWGSPSTHINSIENYQVLEBKMS 379  
 Db 297 FKKEGHSKRGQVYLGVDMEKARAYDLAALKYWGSPSTHINSIENYQVLEBKMS 356  
 Qy 380 ROEYVAHLRRKSSGFSRGSASTYRGVTRHGHQWQARIGVAGNKDLYLGTFQOEBAAE 439  
 Db 357 ROEYVAHLRRKTSGFSRGSASTYRGVTRHGHQWQARIGVAGNKDLYLGTFQOEBAAE 416  
 Qy 440 AYDVAALFKRGANAVTNFDISRVDYERIMASSNLLAGELARRKKNDPRKIDYKSAVY 499  
 Db 417 AYDVAALFKRGANAVTNFDISRVDYERIMASSNLLAGELARRKKNDPRKIDYKSAVY 465  
 Qy 500 TSVNNEE---TVQVQAGNNNNENDESMKVLFNHPSQQQOANGSDQKIMNCGNTRNSA 556  
 Db 466 RNISDEEALTAIVVAGSNKEV-----GSPERVLSF----- 496  
 Qy 557 FSMALQDLIGDVSQGHNMLEDSSKIGTHFSNTSLVTSLSRSRASPBRKGPILFP 616  
 Db 497 -----PTIFA 501  
 Qy 617 MPMEETKI--VNPIGTSVSWLPSPPTVQMRPSPALSHLPVFAWTD 663  
 Db 502 LPQVGPXMGANVVG--NMSWTTNPNADLK--TVSLTLQPMVFAWADS 548

RESULT 5  
 US-08-700-152A-4  
 / Sequence 4, Application US/08700152A  
 / Patent No. 5994622  
 / GENERAL INFORMATION:  
 / APPLICANT: Jofuku, K. Diane  
 / APPLICANT: Okamoto, Jack K.  
 / TITLE OF INVENTION: Methods for Improving Seeds  
 / NUMBER OF SEQUENCES: 4  
 / CORRESPONDENCE ADDRESSES:  
 / ADDRESSER: Townsend and Townsend and Crew LLP  
 / STREET: Two Embarcadero Center, Eighth Floor  
 / CITY: San Francisco  
 / STATE: California  
 / COUNTRY: USA  
 / ZIP: 94111-3834  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Floppy disk  
 / COMPUTER: IBM PC compatible  
 / OPERATING SYSTEM: PC-DOS/MS-DOS  
 / SOFTWARE: PatentIn Release #1.0, Version #1.30  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: US/08/700,152A  
 / FILING DATE: 20-AUG-1996  
 / CLASSIFICATION: 800  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: Bastian, Kevin L.  
 / REGISTRATION NUMBER: 34,774  
 / REFERENCE/DOCKET NUMBER: 023070-067200US  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: (415) 576-0200  
 / TELEFAX: (415) 576-0300



RESULT 7  
 US-09-026-039-9  
 Sequence 9, Application US/09026039  
 Parent No. 638567  
 GENERAL INFORMATION:  
 APPLICANT: Jotiku, K. Diane  
 APPLICANT: Okamuro, Jack K.  
 TITLE OF INVENTION: Methods for Improving Seeds  
 NUMBER OF SEQUENCES: 103  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/026,039  
 FILING DATE: 19-FEB-1998  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/912,272  
 FILING DATE: 15-AUG-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997  
PRIOR APPLICATION DATA: US 08/700,152  
APPLICATION NUMBER: US 08/700,152  
FILING DATE: 20-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Baetian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067230US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 77 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..77  
OTHER INFORMATION: /note= "ANT-R1 direct repeat"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 44..59  
OTHER INFORMATION: /note= "putative ANT-R1 amphipathic  
OTHER INFORMATION: alpha-helix (SEQ ID NO:37)"  
US-09-026-039-9

Query Match 11.0%; Score 385; DB 4; Length 77;  
Best Local Similarity 89.6%; Pred. No. 1e-30; 5; Indels 0; Gaps 0;  
Matches 69; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 296 TSOYRGVTRHMTGRYBAHLMDNSCKEKGOTRKGRQVYLGGMEEKARAYDLALKTW 355  
DB 1 TSOYRGVTRHMTGRYBAHLMDNSCKEKGOTRKGRQVYLGGMEEKARAYDLALKTW 60

QY 356 GPSTHTNFSANVQKI 372  
DB 61 GPSTHTNFSANVQKI 77

RESULT 8  
US-08-912-272-10  
Sequence 10, Application US/08912272  
Patent No. 6093874  
GENERAL INFORMATION:  
APPLICANT: Jofuku, K. Diane  
APPLICANT: Okamura, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,272  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/700,152  
FILING DATE: 20-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Baetian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067220US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 69 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..69  
OTHER INFORMATION: /note= "ANT-R2 direct repeat"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 37..51  
OTHER INFORMATION: /note= "putative ANT-R2 amphipathic  
OTHER INFORMATION: alpha-helix"  
US-08-912-272-10

Query Match 9.9%; Score 346; DB 3; Length 69;  
Best Local Similarity 94.2%; Pred. No. 6.6e-27; 2; Indels 0; Gaps 0;  
Matches 65; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 398 ASYRGVTRHHQSGMQARIGRVAGNKDLYGTFTSEEAAYDAVVAIKRGANAYTNF 457  
DB 1 ASYRGVTRHHQSGMQARIGRVAGNKDLYGTFTSEEAAYDAVVAIKRGANAYTNF 60

QY 458 DISRYDVR 466  
DB 61 DISRYDVR 69

RESULT 9  
US-09-026-039-10  
Sequence 10, Application US/09026039  
Patent No. 6329567  
GENERAL INFORMATION:  
APPLICANT: Jofuku, K. Diane  
APPLICANT: Okamura, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/026,039  
FILING DATE: 19-FEB-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,272  
FILING DATE: 15-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/700,152

FILING DATE: 20-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Baetian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067230US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 69 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..69  
OTHER INFORMATION: /note= "ANT-R2 direct repeat"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 37..51  
OTHER INFORMATION: /note= "putative ANT-R2 amphipathic  
OTHER INFORMATION: alpha-helix"  
US-09-026-039-10

Query Match 9.9% Score 346; DB 4; Length 69;  
Best Local Similarity 94.2% Pred. No. 6.6e-27;  
Matches 65; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 338 ASIVGVTRHGHGHWQARIGVAGNKLVIETSTGEAAEAYDAIKRGANAVNF 457  
DB 1 ASIVGVTRHGHGHWQARIGVAGNKLVIETSTGEAAEAYDAIKRGANAVNF 60

QY 458 DISRYDVR 466  
DB 61 DITRYDVR 69

RESULT 10

US-08-700-152A-1

Sequence 1, Application US/08700152A

Patent No. 5994622

GENERAL INFORMATION:

APPLICANT: Jofuku, K. Diane

TITLE OF INVENTION: Jack K.

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/700,152A

FILING DATE: 20-AUG-1996

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Baetian, Kevin L.

REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 023070-067200US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0300

TELEFAX: (415) 576-0200

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 67 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 1..67  
OTHER INFORMATION: /note= "APETALA 2 (AP2) of Arabidopsis  
OTHER INFORMATION: direct repeat domain AP2-R1 consisting  
OTHER INFORMATION: of amino acids 129-195 of the AP2  
OTHER INFORMATION: protein"  
US-08-700-152A-1

Query Match 5.6% Score 197; DB 2; Length 67;  
Best Local Similarity 51.9% Pred. No. 4.5e-12;  
Matches 40; Conservative 12; Mismatches 15; Indels 10; Gaps 2;

QY 296 TSQYGVTRHRTGTYEAAHMDNSCKEGQTRKGRQVTLGYDMEKAAAYDLAALKY 355  
DB 1 SSQYGVTFYRTGFWESHWD-C-----GRQVTLGFTAAHAAAYDRAAIKFR 50

QY 356 GPSTHINPSEIENYQVL 372  
DB 51 GVEADINFINIDYDDL 67

RESULT 11

US-08-912-272-11

Sequence 11, Application US/08912272

Patent No. 6033874

GENERAL INFORMATION:

APPLICANT: Jofuku, K. Diane

TITLE OF INVENTION: Jack K.

NUMBER OF SEQUENCES: 103

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/912,272

FILING DATE: 15-AUG-1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/879,827

FILING DATE: 20-JUN-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/700,152

FILING DATE: 20-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Baetian, Kevin L.

REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 023070-067220US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0300

TELEFAX: (415) 576-0200

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 67 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Peptide  
LOCATION: 1..67  
OTHER INFORMATION: /note= "RAP2.7-R1 direct repeat"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 33..49  
OTHER INFORMATION: /note= "putative RAP2.7-R1 amphipathic"  
OTHER INFORMATION: alpha helix (SEQ ID NO:36)"  
US-08-912-272-11

Query Match 5.5%; Score 192.5; DB 3; Length 67;  
Best Local Similarity 50.0%; Pred. No. 1.3e-11;  
Matches 40; Conservative 14; Mismatches 13; Indels 13; Gaps 3;  
CY 296 TSGYRGVTRHWTGRYEAHLMDNSCKKEGQTRKRGQVYLGVDMEKAAAYDLAALKYW 355  
DB 1 SSGYRGVTFYRRTGFWESH:WD-C-----GKQVYLGSGFTAAHAAAYDRAAIKFR 50  
CY 356 GPSTHINFSEIENYQVLEEM 375  
DB 51 GVDADINFTLGDIYE--EDM 67

RESULT 12  
US-09-026-039-11  
Sequence 11, Application US/09026039  
Patent No. 6329567  
GENERAL INFORMATION:  
APPLICANT: Jofuku, K. Diane  
APPLICANT: Okamoto, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/026,039  
FILING DATE: 19-FEB-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,272  
FILING DATE: 15-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/700,152  
FILING DATE: 20-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067230US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 67 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide

LOCATION: 1..67  
OTHER INFORMATION: /note= "RAP2.7-R1 direct repeat"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 33..49  
OTHER INFORMATION: /note= "putative RAP2.7-R1 amphipathic"  
OTHER INFORMATION: alpha helix (SEQ ID NO:36)"  
US-09-026-039-11

Query Match 5.5%; Score 192.5; DB 4; Length 67;  
Best Local Similarity 50.0%; Pred. No. 1.3e-11;  
Matches 40; Conservative 14; Mismatches 13; Indels 13; Gaps 3;  
CY 296 TSGYRGVTRHWTGRYEAHLMDNSCKKEGQTRKRGQVYLGVDMEKAAAYDLAALKYW 355  
DB 1 SSGYRGVTFYRRTGFWESH:WD-C-----GKQVYLGSGFTAAHAAAYDRAAIKFR 50  
CY 356 GPSTHINFSEIENYQVLEEM 375  
DB 51 GVDADINFTLGDIYE--EDM 67

RESULT 13  
US-08-700-152A-2  
Sequence 2, Application US/08700152A  
Patent No. 5994692  
GENERAL INFORMATION:  
APPLICANT: Jofuku, K. Diane  
APPLICANT: Okamoto, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/700,152A  
FILING DATE: 20-AUG-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067200US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 68 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 1..68  
OTHER INFORMATION: /note= "APETALA 2 (AP2) of Arabidopsis  
OTHER INFORMATION: direct repeat domain AP2-R2 consisting  
OTHER INFORMATION: of amino acids 221-288 of the AP2  
OTHER INFORMATION: protein"  
US-08-700-152A-2

Query Match 5.5%; Score 191.5; DB 2; Length 68;  
Best Local Similarity 60.3%; Pred. No. 1.6e-11;  
Matches 41; Conservative 8; Mismatches 16; Indels 1; Gaps 1;

QY 398 ASIVRGVTRHHQHGWRQAR-GRVAGNKDLYLTGFTSTQBEAAEYDVAIKFRGANAVTNF 457  
1 SSKYRGVTL-HKCGRWEARMGQFLGKKYVYLGFDTEVEAARAVDAIKKNGKDAVTNF 59  
DB 458 DISRYDVE 465  
60 DPSIYDEE 67

RESULT 14  
US-08-912-272-5  
Sequence 5, Application US/08912272  
Patent No. 603874  
GENERAL INFORMATION:  
APPLICANT: Jotuku, K. Diane  
APPLICANT: Okamoto, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,272  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/700,152  
FILING DATE: 20-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067220US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 68 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..68  
OTHER INFORMATION: /note="AP2-R2 direct repeat at  
OTHER INFORMATION: positions 221 to 288"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 33..50  
OTHER INFORMATION: /note="putative AP2-R2 amphipathic  
OTHER INFORMATION: alpha-helix (SEQ ID NO:7)"  
US-08-912-272-5

Query Match 5.5%; Score 191.5; DB 3; Length 68;  
Best Local Similarity 60.3%; Pred. No. 1.6e-11;  
Matches 41; Conservative 8; Mismatches 18; Indels 1; Gaps 1;  
QY 398 ASIVRGVTRHHQHGWRQARIGRVAGNKDLYLTGFTSTQBEAAEYDVAIKFRGANAVTNF 457  
1 SSKYRGVTL-HKCGRWEARMGQFLGKKYVYLGFDTEVEAARAVDAIKKNGKDAVTNF 59

DB 1 SSKYRGVTL-HKCGRWEARMGQFLGKKYVYLGFDTEVEAARAVDAIKKNGKDAVTNF 59  
QY 458 DISRYDVE 465  
DB 60 DPSIYDEE 67

RESULT 15  
US-09-026-039-5  
Sequence 5, Application US/09026039  
Patent No. 6329567  
GENERAL INFORMATION:  
APPLICANT: Jotuku, K. Diane  
APPLICANT: Okamoto, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/026,039  
FILING DATE: 19-FEB-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,272  
FILING DATE: 15-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/700,152  
FILING DATE: 20-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067230US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 68 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..68  
OTHER INFORMATION: /note="AP2-R2 direct repeat at  
OTHER INFORMATION: positions 221 to 288"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 33..50  
OTHER INFORMATION: /note="putative AP2-R2 amphipathic  
OTHER INFORMATION: alpha-helix (SEQ ID NO:7)"  
US-09-026-039-5

Query Match 5.5%; Score 191.5; DB 4; Length 68;  
Best Local Similarity 60.3%; Pred. No. 1.6e-11;  
Matches 41; Conservative 8; Mismatches 18; Indels 1; Gaps 1;  
QY 398 ASIVRGVTRHHQHGWRQARIGRVAGNKDLYLTGFTSTQBEAAEYDVAIKFRGANAVTNF 457  
1 SSKYRGVTL-HKCGRWEARMGQFLGKKYVYLGFDTEVEAARAVDAIKKNGKDAVTNF 59

Wed Mar 17 08:18:55 2004

us-10-024-632-2.rat

Page 9

Db 1 SSKYGVTL-HKCGWEARNGQFLGKXVYIGLFDTEVEARARAYDKAIAIKONGDAVTNF 59  
QY 458 DISRYDVE 465  
Db 60 DPSYDDE 67

Search completed: March 9, 2004, 10:48:10  
Job time : 25 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 13, 2004, 00:46:03 ; Search time 98 Seconds  
(without alignments)

3754.413 Million cell updates/sec

Title: US-10-024-632-2

Perfect score: 3502

Sequence: 1 MKRINSSNNTDDGNHNLG.....RSPALSLHLPVPSWTDT 663

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO/US10024632/runat\_09032004\_101208\_10797/app\_query.fasta\_1.839  
-DB=Issued\_Patents\_NA -QEMT=fastap -SUFFIX=rnt -MINMATCH=0.1 -LOOFCU=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=numan40.cdt  
-LIST=45 -DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10024632@cgn2\_1.1.44@runat\_09032004\_101208\_10797 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG SCORES=0 -MATT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MATT\_TIMEOUT=30 -THRAIDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiltest.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1338	37.9	2148	4	US-09-227-421-1
2	1328	37.9	2148	4	US-09-479-855-1
3	1307	37.3	1738	4	US-09-227-421-4
4	1307	37.3	1738	4	US-09-479-855-4
5	517	14.8	1680	2	US-08-700-152A-3
6	292	8.3	11721	4	US-09-026-039-3
7	292	8.3	11721	4	US-09-026-039-3
8	263	7.5	101	4	US-09-113-294A-6727
9	158.5	4.5	888	4	US-09-198-119C-64
10	156.5	4.5	1405	4	US-09-202-161B-6
11	154.5	4.4	1669	4	US-09-026-039-1
12	148.5	4.2	854	4	US-09-198-119C-86

13	148	4.2	950	4	US-09-198-119C-70	Sequence 70, Appl
14	147	4.2	874	4	US-09-198-119C-52	Sequence 52, Appl
15	147	4.2	1055	4	US-09-533-029-21	Sequence 21, Appl
16	145.5	4.2	1577	4	US-09-533-029-11	Sequence 11, Appl
17	142	4.1	294	4	US-09-313-284A-6087	Sequence 6087, Ap
18	141.5	4.0	1281	4	US-09-533-029-65	Sequence 65, Appl
19	140.5	4.0	751	4	US-09-533-029-81	Sequence 81, Appl
20	140	4.0	1500	3	US-09-300-672-1	Sequence 1, Appl
21	139.5	4.0	832	4	US-09-198-119C-82	Sequence 82, Appl
22	139	4.0	8146	4	US-09-976-594-725	Sequence 725, App
23	139	4.0	11917	4	US-09-866-921-32	Sequence 32, Appl
24	137.5	3.9	913	4	US-09-533-029-33	Sequence 33, Appl
25	137	3.9	1499	3	US-09-300-672-3	Sequence 3, Appl
26	136	3.9	632	4	US-09-198-119C-18	Sequence 18, Appl
27	136	3.9	632	4	US-09-198-119C-46	Sequence 46, Appl
28	135	3.9	933	4	US-09-202-161B-4	Sequence 4, Appl
29	134.5	3.8	174	4	US-09-202-161B-21	Sequence 21, Appl
30	134.5	3.8	682	4	US-09-198-119C-92	Sequence 92, Appl
31	134.5	3.8	884	4	US-09-198-119C-50	Sequence 50, Appl
32	133.5	3.8	830	4	US-09-198-119C-84	Sequence 84, Appl
33	133.5	3.8	884	4	US-09-198-119C-76	Sequence 76, Appl
34	132.5	3.8	887	4	US-09-198-119C-68	Sequence 68, Appl
35	132.5	3.8	889	4	US-09-198-119C-62	Sequence 62, Appl
36	132.5	3.8	953	4	US-09-198-119C-80	Sequence 80, Appl
37	132	3.8	793	4	US-09-198-119C-80	Sequence 90, Appl
38	131.5	3.8	1132	3	US-08-894-731-3	Sequence 3, Appl
39	131	3.7	769	4	US-09-198-119C-58	Sequence 58, Appl
40	130.5	3.7	785	4	US-09-198-119C-80	Sequence 80, Appl
41	130.5	3.7	806	4	US-09-198-119C-78	Sequence 78, Appl
42	130.5	3.7	876	4	US-09-198-119C-48	Sequence 48, Appl
43	130.5	3.7	1132	4	US-09-198-119C-56	Sequence 56, Appl
44	130	3.7	937	4	US-09-301-666A-5	Sequence 5, Appl
45	130	3.7	937	4	US-09-301-217-5	Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-09-227-421-1  
Sequence 1, Application US/09227421  
Patent No. 6559357  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Mizukami, Yukiko  
TITLE OF INVENTION: The Regents of the University of California  
TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility  
FILE REFERENCE: 023070-090700PC  
CURRENT APPLICATION NUMBER: US/09/227,421  
PRIOR FILING DATE: 1999-01-08  
PRIOR APPLICATION NUMBER: US-09/227,421  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 2148  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (269)..(1936)  
OTHER INFORMATION: AINTEGUMENTA (ANT)  
NAME/KEY: misc feature  
LOCATION: (1105)..(1339)  
OTHER INFORMATION: encodes first AP2 domain  
NAME/KEY: misc feature  
LOCATION: (1340)..(1414)  
OTHER INFORMATION: encodes linker region  
NAME/KEY: misc feature  
LOCATION: (1415)..(1621)  
OTHER INFORMATION: encodes second AP2 domain  
US-09-227-421-1

Alignment Scores:  
 Pred. No.: 3,44e-130 Length: 2148  
 Score: 1328.00 Matches: 312  
 Percent Similarity: 52.84% Conservative: 69  
 Best Local Similarity: 43.27% Mismatches: 116  
 Query Match: 37.92% Indels: 224  
 Gaps: 21  
 DB: 4

US-10-024-632-2 (1-663) x US-09-227-421-1 (1-2148)

QY 1 MetLysArgLLeuGlnSerAsnThrAspGlyAsnAsnHisLeuTyrLeuGly 20  
 DB 269 ATGAAGCTCTTTTGGATGAATGATATATATCATACCAACGACATTAATTTGTATGG 328  
 QY 21 PheSerLeuSerProHis-MetLysMet----- 29  
 DB 329 TTCTCATTTCTCTTCAATATGATGAAATGGAGGTAGAGAGGATAGAGCTATTATAC 388  
 QY 30 -----GluAlaThrSerAlaAlaThrValProThrThrPheTyrMet 43  
 DB 389 TCATCTTCAACTCTTCTCACTGCAACTCTTCTTCTGTTCCACTCAACTGTTGTT 448  
 QY 44 SerProSerGlnSerHisLeuSerAsnPheGlyMetCysTyrGlyValGlyLysGly 63  
 DB 449 GGT-----GACAACTGCTGCAACTTGGTGTGTTGCTATGATCAACCAATGGA 499  
 QY 64 AsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeuCysLleLeu 83  
 DB 500 GGAATCTATTTCTACATGCTGTGATGACCACTCAATCTGATGTTCTTCTGCTATG 559  
 QY 84 GluAlaLeuLysArgSerGlnThrGlnValMetValProThrSerSerProLysLeuGly 103  
 DB 560 GAAGCTCTCAACAGATCTTCTCACTGCAATCCATCAAGATCTCTCCAAAGCTGAG 619  
 QY 104 AspPheLeuGlyGlyAlaThrMetGlyThrHisGlyTyrGlySerHisGlnArgGly-- 122  
 DB 620 GATTTCTTTGGG-----ACCATCAACAACAACAGTCAACAAGAACCCATG 667  
 QY 123 ---LeuSerLeuAspSerLleTyrTyrAsnSerGlnAsnAlaGluAlaGlyProAsnArg 141  
 DB 668 GATCTAGCTTGAATGTTATTTATTTCAACACCACTCAT-----GAGCCCAACG 718  
 QY 142 AspLeuLeuSerGlnProPhe-----ArgGlnGlnGlyHisMetSerValGln 157  
 DB 719 ACTACAACTTCAAGAGTTCTTAGCTCCCTCAACCAACAGAACCAT----- 766  
 QY 158 ThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGlnGly 177  
 DB 767 -----GAGGAA 772  
 QY 178 GluThrThrLysGlnThrHisValSerAspCysSerSerLeuMetProGlnMetThrGln 197  
 DB 773 GAAACT---AGAAATTACGGGATGAC-----CCATGTTTGAACACT 811  
 QY 198 GlyLeuLysAsnTyrValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGly 217  
 DB 812 GGA----- 814  
 QY 218 GlnGlnMetAsnCysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGly 237  
 DB 815 GGATCTTTTATATAGGAGATATGGGAATTCAA----- 850  
 QY 238 CysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerCysVal 257  
 DB 851 -----CAGTCACTGAGCTTATCATGAGCCCTGGGTCACAAATCTAGCTGCATC 898  
 QY 258 ThrAlaProSer----- 261  
 DB 899 ACTGGCTCTACCAACCAACCAAAACCAAAACCAACCAACCAACCAACG 958  
 QY 262 -----GlyThrAspSerValAlaVal 268  
 DB 959 CAGATCTGTAGAGCTTTGTGAGACAGAGCTGGGTTTGAAGACGACAAATGGGGCT 1018

QY 269 AspAlaLysLysArgGlyHis-----AlaLysLeuGlyGlnLysGlnProValHis 285  
 DB 1019 GCGAAGAGAGAGAGGAGCAAGAGAGATGTTGATGTTGGTCAAGAAACAGATGTTTCAT 1078  
 QY 286 ArgLysSerLleAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValThrArgHis 305  
 DB 1079 AGAAATCTATGATACCTTTTGGACACGAACTTCTCAATACCGAGGGCTTCAAGACAT 1138  
 QY 306 ArgTyrThrGlyArgTyrGlnAlaHisLeuTyrAspAsnSerCysLysLysGlnGln 325  
 DB 1139 AGATGAGCTGATGATATATGATAGCTCATCTATGACCAATAGTTTCAAGAGAGAGTCAAC 1198  
 QY 326 ThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnGlnLysAlaAlaArg 345  
 DB 1199 AGTAAAGAAAGAAAGAAAGATTTATCTGGAGGTTATGATATGAGAGAGAACTGCTGCA 1258  
 QY 346 AlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyProSerThrHisLleAsnPheSerLle 365  
 DB 1259 GCATATGATCTTGTGCACTCAAGTATCTGGGTCCTTCACTCAACCAATTTCTTGGC 1318  
 QY 366 GluAsnTyrGlnValGlnLeuGlnLysMetLysAsnMetSerArgGlnGlnTyrValAla 385  
 DB 1319 GAGAAATTAACGAAAGAGATTTGAGACATGAAAGAACATCACTAGACAAAGATATGTGCA 1378  
 QY 386 HisLeuArgArgLysSerSerGlyPheSerArgLysLleTyrArgGlyValThr 405  
 DB 1379 CATTTGAGAAAGAAAGACAGTGTCTCTAGGGGTCTTCATCATATGAGAGGTCA 1438  
 QY 406 ArgHisHisGlnHisGlyArgTyrGlnAlaArgLleGlyArgValAlaGlyAsnLysAsp 425  
 DB 1439 AGACATCACACAGCATGAAAGTGGCAAGCAGGATGTAGAGTGTGAGTGAACAAAGAT 1498  
 QY 426 LeuTyrLeuGlyThrPheSerThrGlnGlnGlnAlaAlaGluAlaTyrAspValAlaAla 445  
 DB 1499 CTCTACCTTGAACCTTTTGGAAACCAAGAAAGCTGCAAGACCTTACATGATAGACGA 1558  
 QY 446 IleLysPheArgGlyAlaAsnAlaValThrAsnPheAspLleSerArgTyrAspValGlu 465  
 DB 1559 ATTAAGTTCGTGGCAAAATGCTGTGACTTATGATATGACGAGGATGATGAT 1618  
 QY 466 ArgLleMetLleAspSerAsnLeuAlaGlyGlnLeuAlaArgArgLysLysAspAsn 485  
 DB 1619 CGTATCATGTCTAGAACACACTCTGTCTGGAGATTTAGCCGCAAGAAC----- 1669  
 QY 486 AspProArgAsnLysAspLleAspTyrAsnLysSerValValThrSerValAsnAsnGln 505  
 DB 1670 -----AACACAGCATTTGTC-----GTACAGAGATTA 1696  
 QY 506 GluThrValGlnValGlnAlaGlyAsnAsnAsnAsnGluAspSerGluTyrLysMet 525  
 DB 1697 GAA----- 1699  
 QY 526 ValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLys 545  
 DB 1700 -----GACCAAC 1708  
 QY 546 IleMetAsnCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuLle 565  
 DB 1709 GCTCTAAATGCT----- 1720  
 QY 566 GlyLleAspSerValGlySerGlyGlnHisAsnMetLeuAspLysSerLysLleGly 585  
 DB 1720 ----- 1720  
 QY 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerArgGluAlaSer 605  
 DB 1721 -----GTTGTGAAGGTGTTTCAACAAAGAGTCA 1753  
 QY 606 ProGluLysArgGlyProSerLeu-----LeuPheProMetProProMetGluThrLys 623  
 DB 1754 ACTCCGAGAGACCTCTTGAATTTTCCGGGCAATTTCCCGTGTGCTCAAGTATATCAAAAG 1813

D	b		620	GATTTCTTGGG-----	ACCACCAACAACAACAAGTCACAAGAAGCATG	667
O	y		123	--LeuSerLeuAspSerIleuTyTrpAsnSerGlnAsnIagiuaIagiInProAsnArg	141	
D	b		668	GATCTTAGCTTGAGTAAGTTATTATCTACACACCACTCAT	718	
O	y		142	AspLeuLeuSerGlnProPhe----	ArgIngInIynHisMetSerValGln	157
D	b		719	ACTACAAACATTCAAAGAGTTCTTAGCTTCCCTCAACACGAAGAACAT	766	
O	y		158	ThriHieProTyTrpTyrSerCylLeuAlaIacyShIsGIeuTyTrpGlnAlaProLeuGlu	177	
D	b		767	-----	GAAGAA	772
O	y		178	GluThrThryLysGluThrHiseValserApsCysserSerLeuMetProGlnMetThrGlu	197	
D	b		773	GAATCT-----	AGAAATTACCGGAATGAC	811
O	y		198	GlyLeuLysAsnTrpValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlu	217	
D	b		812	GGA-----		814
O	y		218	GlnGlnMetAsnCysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGly	237	
D	b		815	GGGTCTTTTAAITGTAAGGGTATAATGGGAATTTCCA	850	
O	y		238	CysGlyGluLeuGlnInsLeuSerLeuSerMetSerProGlySerGlnSerCysVal	257	
D	b		851	-----	CAGTCACTGAGCTTATCCATGAGCCCTGGGTCAATCTAGCTGCATC	898
O	y		258	ThraIaProSer-----		261
D	b		899	ACTGGCTTCACCAACCAACCAAAAACCAAAACCAACGCCAAAAACCAACAG	958	
O	y		262	-----	GlyThrAspSerValAlaVal	268
D	b		959	CAGATCTCTGAAGCTCTTGCGAGACAAGCGTTGGTTGAGACGACACATGGCGCT	101	
O	y		269	AspAlaIalysIarGlyHis-----	IalysIeugIyGlnIyVgInProValHis	285
D	b		1019	GCGAAGAAAGAGGGGACAAGAAGATGTTGTAGTTGGTGCMAAACAAGATTGTTCAT	107	
O	y		286	ArgIysSerIleAspThrPheGlyGlnArgThrseryInTyFargIyValThrArgHis	305	
D	b		1079	AGAAATCTATGCATCTTTGGACACAGCACTTCTCAATACCGAGGGGTACAAGACAT	113	
O	y		306	ArgTrpThrgIyArgTyTrgIuAlaHisIeunTPapaAsnSerCysIyIyGlnGlyGln	322	
D	b		1139	AGATGGACTGTGATATGAGTCACTCATGGGCAATAGTTTCMAAAGAAAGGTAC	119	
O	y		326	ThurArgIySglIyArgGlnValTyTrLeuGlyGlyTyRasPmetGluGluValAlaArg	345	
D	b		1199	AGTGAAGAAAGAGACAAAGTTATCTGGAGGTTTGTATGATGAGAGCAAGAGCTGCTGA	122	
O	y		346	AlaTyRasPleuAlaIaleuLySTyTrPGlyProSerThrHisIleAsnPheserIle	365	
D	b		1259	GCAATATATCTTCTGTGACTCAAGTACTGGGGTCCCTCTACTCACACCAATTTCTGTGCG	133	
O	y		366	GluunTyTrgInValGlnLeuGluGluMetLysAsnMetSerArgGlnGluTyTrValAla	385	
D	b		1319	GAGAAITHTCAGAAAGATGATGAGACATGAAGACATGACTAGACAAAGATATGTTGCA	133	
O	y		386	HisLeuArgIyLysSerSerGlyPheSerArgGlyAlaserIleTyFargIyValThr	405	
D	b		1379	CATTGGAGAGGAAGACAGCTGTTCTCTTAAGGGGTCTTCATCTTAAAGGAGATCCA	144	
O	y		406	ArghSHIsGlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysASP	422	
D	b		1439	AGACATCCACAGCTGGAAGGTGGCAAGACGAGTTGTAGGTGCGCTGGAAACCAAGAT	144	
O	y		426	LeuTyTrLeuGlyThrpheSerThrgInGluGluAlaIagluAlaTyRasPvalAlaIa	444	

```

Db 1499 CTCTACCTTGGAACCTTTGGAAACCAAGAAAGCTGCAAGCTTACGATGTCAGACGA 1558
Qy 446 11lelyspheargglyalaasmlaavalThrAspPheAspIleSerArgTyrAspValGlu 465
1559 ATTAAAGTTCCGTGCGACAAATGCTGATCTACTATCTCATATCAGAGGTGATGTTGAT 1618
Qy 466 ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGluLeuAlaArgArgLysLysAspAsn 485
1619 CGATTCATGCTGTAGTAACACACTCTTCTGTGAGAGAGTACCGCAAGAAC----- 1669
Qy 486 AspProArgAsnLysAspIleAspTyrAsnLysSerValValThrSerValAsnAsnGlu 505
1670 -----AACAAACGACTTGTGC-----GTCAAGAAATAC 1696
Qy 506 GluThrValGluValGluAlaGlyAsnAsnAsnAsnGluAsnAspSerGluTyrLysMet 525
1697 GAA----- 1699
Qy 526 ValLeuPheAsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLys 545
1700 -----GACCAAAAC 1708
Qy 546 11leucaenCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565
1709 GCTCTAAATGCT----- 1720
Qy 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerLysIleGly 585
1720 ----- 1720
Qy 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGluAlaSer 605
1721 -----GTGTGGAAGGTGTCTCCAAACAAAGAGTCAGAT 1753
Qy 606 ProGluLysArgGlyProSerLeu-----LeuPheProMetProProMetGluThrLys 623
1754 ACTCCCGAGAGACTCTGAGTTTCCGCGCGATTTTCGCGTTCCTCAAGATTAATCAAAG 1813
Qy 624 11leval-----AspProIleGlyThrSerValThrSerTyrLeuProSerProThrValGln 642
1814 ATGTTCCGATCAATATGCGCGAAATATAGTCTTCTCAAGATCAAACTTAATGCTGAG 1873
Qy 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTyrThrAsp 662
1874 CTTAAG--ACCGTCGCTCTTAATCTTGCCTCAGATGCCGTTTTCGCTGCTGGCGTGA 1930
Qy 663 Thr 663
1931 TCT 1933

```

RESULT 3  
US-09-227-421-4  
Sequence 4, Application US/09227421

Patent No. 6559357  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Mizukami, Yukiko  
TITLE OF INVENTION: The Regents of the University of California  
TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility  
FILE REFERENCE: 023070-090700PC  
CURRENT FILING DATE: US/09/227,421  
PRIOR APPLICATION NUMBER: US 09/227,421  
PRIOR FILING DATE: 1999-01-08  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 4  
LENGTH: 1738  
TYPE: DNA  
ORGANISM: Brassica napus  
NAME/KEY: CDS

```

; LOCATION: (1) ..(1647)
; OTHER INFORMATION: BANT, ANT ortholog from Brassica napus (canola)
US-09-227-421-4
Alignment Scores:
Pred. No.: 4,01e-128 Length: 1738
Score: 1307.00 Matches: 308
Percent Similarity: 54.52% Conservative: 78
Best Local Similarity: 43.50% Mismatches: 104
Query Match: 37.32% Indels: 218
DB: 4 Gaps: 25
US-10-024-632-2 (1-663) x US-09-227-421-4 (1-1738)
Qy 9 AsnThrAspAspArgLysAsnAsnHisAsnTyrLeuGlyPheSerLeuSerProHisMet--- 27
16 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 75
Qy 28 -----LysMetGluAlaThrSerAlaAlaThrValProThrThrPhe 41
76 AAATGGGTGGTGGAGAGCTCTTACTCAGCTTCGTCCTTCAAGTTCCTCTCTCTCTCTCT 132
Qy 42 TyrMetSerProSerGln-----SerHisLeuSerAsnProMetGlyMetCysTyr 57
133 ---TCTGTCCACCAACACACTTGTGTGGCAACACAGTACATATGAGTTGCTAC 189
Qy 58 GlyValGly---GluAsnGlyAsnPheHisSerProLeuThrValMetProLeuLysSer 76
190 GGTTCATTAAGTACAGCAGTAAAGGAATGATATCTCAATAGTGTGATGATGATGATGAT 249
Qy 77 AspGlySerLeuCysIleLeuGluAlaLeuLysArgSer----- 89
250 GACGGTTCCTCTTGTATGAGAACTCTCAACAGATCTTCTCACTCGAATATATATAC 309
Qy 90 GlnThrGlnValMetValProThrSerSerProLysLeuGluAspPheLeuGlyGlyAla 109
310 CATATCAAGTT-----TCATCTCAAAAGATGAGAAATTTT----- 345
Qy 110 ThrMetGlyThrHisGluTyrGly---SerHisGluArgGly-----LeuSerLeuAsp 126
346 ---TTTGGAGCCCATCATCATACACACACACACACACACACACACACACACACAT 402
Qy 127 SerIleTyrAsnSerGlnAsnAlaGluAlaGlnProAsnArgAspLeuLeuSerGln 146
403 AGTTATTTCTACAAATACATCATCATGCG-----CCAAACAAACAAACCACTTTCAA 453
Qy 147 ProPheArgGlnGlnGlnGlyHisMetSerValGlnThrHisProTyrTyrSerGlyLeuAla 166
454 GAGTTC-----TTTACCTTCCCTCAAACTAGAAACCAACAT----- 489
Qy 167 CysHisGlyLeuTyrGlnAlaProLeuGlnGlnGluTyrThrLysGluThrHisValSer 186
490 -----GAGGAAGAAACA-----AGAAACTACAGAAAT 516
Qy 187 AspCysSerSerLeuMetProGlnMetThrGlnGlyLeuLysAsnTyrValAlaProThr 206
517 GAC-----CTTGTTTGACACATGGA----- 537
Qy 207 ArgGluPheSerThrHisGlnGlnValLeuGlnGlnGlnMetAsnCysGlyMetCysLysn 226
538 -----GAGAGGCTTTTAATGTAAGGGGTATATGCG 567
Qy 227 GluArgAsnGlyValSerLeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeu 246
568 GAATTTCA-----CAATCACTAGAGCTTG 591
Qy 247 SerMetSerProGlySerGlnSerSerCysValThrAla----- 259
592 TCCATGAGCCCTGGGTCAAACTAGCTGATCACTGCTTCATCACACCAAAACCAA 651
Qy 260 -----ProSer 261
652 ACTCAAAACCAACAGCAGATCTCTGAAGCTTTGTGAGACAAAGTGTGATTTGAGACA 711

```





NAME: Basfian, Kevin L.  
 REGISTRATION NUMBER: 34,774  
 REFERENCE/DOCKET NUMBER: 023070-067200US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1680 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 190..1488  
 OTHER INFORMATION: /product= "APETALA 2 (AP2)"  
 US-08-700-152A-3

Alignment Scores:  
 Pred. No.: 1.58e-44 Length: 1680  
 Score: 517.00 Matches: 163  
 Percent Similarity: 43.76% Conservative: 65  
 Best Local Similarity: 31.29% Mismatches: 151  
 Query Match: 14.76% Indels: 142  
 DB: 2 Gaps: 19

US-10-024-632-2 (1-663) x US-08-700-152A-3 (1-1680)

```

QY 171 TyrglnAlaProLeuGluGluGluThrThryseGluThrHis-----184
DB 214 CACCAACAACAAGAGAGAGAAATCTAAGAGATTGTTATTTCTTACCAAGTAACGG 273
QY 185 -----ValSerAspCysSerSerLeuMetProGlnMetThrGluGlyLeuLys---200
DB 214 GTTGATCTTTCTCTTAATTCAGCTCTTCACTGTGTGTATCGAAGATGGATCCGATGAC 333
QY 201 -----AenTPValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlu 217
DB 334 GATCAACTTAACCGGCTGACACCAATTAACCACTGTGTACCCATCGATTCTTCCCTGAG 393
QY 218 GlnGlnMetAsnCysGlyMetGlyAsnGluArgAsnGlyValSer-----232
DB 394 ATGATTTCTTAACGGCGGT-----GGTGTCTCTTCGCGCTTCCCTCGG 435
QY 233 -----LeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeuSerMetSer 249
DB 436 GCTCACTGGTTGGTGTAAAGTTTGT-----462
QY 250 ProGlySerGlnSerSerCysValThrAlaProSerGlyThrAspSerValAlaValAsp 269
DB 463 -----CAGTCCGATCTAGCCACCGGATGTCGCGGGGTAAAGCTAACCAAGTTGCC 513
QY 270 AlaIysLysArgGlyHisAlaLysLeuGlyGlnGlnProValHisArgLysSerIle 289
DB 514 GCT-----GCCGTAGTGAGCGCGGACAGCCGTTG---AAAAAGAGTCGG 555
QY 290 AspThrPheGlyGlnArgThrSerGlnThrArgGlyValThrArgHisArgThrThyGly 309
DB 556 CGTGACCAAGATCAAGAGATTCAGATTAAGAGGTGATGATTACGTCGCGGATCCGGA 615
QY 310 ArgTyrglnAlaHisLeuThrAspAsnSerCysLysLysGlnGlyGlnThrArgLysGly 329
DB 616 AAGTGGGAATCTCAATTTGGGAC-----TGT-----GGG 645
QY 330 ArgGlnValTyrglnGlyGlyTyrglnPheMetGluGluValAlaAlaArgAlaTyrgln 349
DB 646 AAACAGATTACTAGTGGATTTGACCTGCTCATGACGCTCGAGCTCGACATGATGATGA 705
QY 350 AlaAlaLeuLysTyrglnPheProSerThrHisIleAsnPheSerIleGluAsnTyrgln 369
DB 706 GCTGCTATTAAATTCGCTGAGATGAAGCGGATATCAATTTCAACATGACGATGATGAT 765

```

```

QY 370 ValGlnLeuGluGluMetLysAsnMetSerArgGlnGluTyrglnAlaHisLeuArgArg 389
DB 766 GATGACTTGAACAAGATGATTAATTTAACCAAGAGAGTTCGTACACGTACTTCGCCGA 825
QY 390 LysSerSerGlyPheSerArgGlyAlaSerIleTyrglnArgGlyValThrArgHisHisGln 409
DB 826 CAAGACAGAGCTTCCCTCGAGAGATTCGAGATTAAGAGGTGATGATCTTGG---CATAG 882
QY 410 HisGlyArgTPGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrglnGly 429
DB 883 TGTGCTCTTGGGAAGCTCGTATGGCTCATATTTCTAGCAAAAAGTATGATTTATTTGGCT 942
QY 430 ThrPheSerThrGlnGluGlnAlaAlaGluAlaTyrglnPheValAlaAlaIleLysPheArg 449
DB 943 TTGTCGACACCGAGCTGAGAGCTGCTGAGCTTACGATTAAGATGCAATCAATGATTAAC 1002
QY 450 GlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrglnPheValGluArgIleMetAla 469
DB 1003 GGCAGAAAGCCGCTGACCACTTGTATCCGATATTTCATGAGAGAACTCAATGCCGAG 1062
QY 470 SerSerAsnLeuLeuAlaGlyGluLeuAlaArgArgLysLysAspAsnAspProArgAsn 489
DB 1063 TCATCAGCG-----AATCTTACTTCTCCACAAT 1092
QY 490 LysAspIleAspTyrglnLysSerValAlaThrSerValAsnAsnGluGluThrValGln 509
DB 1093 CACAACTCGAT-----1104
QY 510 ValGlnAlaGlyAsnAsnAsnGlnLysAsnSerGluTyrglnMetValLeuPheAsn 529
DB 1105 CTGAGCTTGGAAATTCGCTAATTCG-----AAG 1134
QY 530 HisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLysIleMetAsnGly 549
DB 1135 CATAAAGTCAAGATATGCGGCTCAGATGAACCAACAACAGAT-----1182
QY 550 GlyAsnTyrglnAsnSerAlaPheSerMetAlaLeuGlnAspLeuIleGlyIleAspSer 569
DB 1183 -----TCTCCACTGTAATGAAGTTCTTGATTAAGTCAA 1218
QY 570 ValGlySerGlyGlnHis-----AsnMetLeuAspLysSerIle 584
DB 1219 ACCGGAATGCTTAACCATTCCTCCATTCMAACCAACCAATTTCCGGGCGACGACACAT 1278
QY 585 GlyThr-----HisPheSer-----589
DB 1279 GGTAGCGAGCGGATTTCTCACTGTTCCGGCGCTGAGAACACCGGTTGATGGTCGG 1338
QY 590 -----AenThrSerSerLeuValThrSerLeuSerSerSerArgGluAlaSerProGluLys 608
DB 1339 GCCTCGACCAACCAAGTGTGACAAATGCTGCAGATCATCAGATTTCTTCTCTCATCAT 1398
QY 609 ArgGlyProSerLeuLeuPheProMetProPheGluThrLysIleValAsnProIle 628
DB 1399 -----CACAAACAGATTTTAATTTACT 1422
QY 629 GlyThrSerValThrSerThrLeuProSerProThrValGln-----MetArgPro 645
DB 1423 TCTACTCTTCATCAAAATGCTGCGACAAATGCTTCCAACTCTCTCTCATGAGACCT 1482
QY 646 Ser 646
DB 1483 TCT 1485

```

RESULT 6  
 US-09-026-039-3  
 Sequence 3, Application US/09026039  
 Patent No. 6328567  
 GENERAL INFORMATION:  
 APPLICANT: Jofuku, K. Diane  
 APPLICANT: Okamoto, Jack K.  
 TITLE OF INVENTION: Methods for Improving Seeds  
 NUMBER OF SEQUENCES: 103



CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/026,039  
 FILING DATE: 19-FEB-1998  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 06/912,272  
 FILING DATE: 15-AUG-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 06/879,827  
 FILING DATE: 20-JUN-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/700,152  
 FILING DATE: 20-AUG-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bastian, Kevin L.  
 REGISTRATION NUMBER: 34,774  
 REFERENCE/DOCKET NUMBER: 023070-067230US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 11721 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: -  
 LOCATION: 1..11721 /note="Arabidopsis ABETL2 (AP2)"  
 OTHER INFORMATION:  
 OTHER INFORMATION: genomic sequence"  
 US-09-026-039-3  
 Alignment Scores:  
 Pred. No.: 2,86e-19 Length: 11721  
 Score: 292.00 Matches: 139  
 Percent Similarity: 34.05% Conservative: 50  
 Best Local Similarity: 25.05% Mismatches: 121  
 Query Match: 8.34% Indels: 246  
 DB: 4 Gaps: 16  
 US-10-024-632-2 (1-663) x US-09-026-039-3 (1-11721)  
 QY 130 TTAAsSerGlnAsnAlaGlnProAsnArgAspLeuLeuSerGlnProPheAr 149  
 Db 1004 TATTAGCTTCTAATCTGAGAGATCATACAGAGATTAATTTGAACCTTCAAG 1063  
 QY 149 GGGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGGNGG 169  
 Db 1064 ATCAAAATCAAGAAACCAAAACCAAAACCAAAACCAAAATGATGGAT 1106  
 QY 169 YLeuTYGlnAlaProLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 184  
 Db 1107 -CTAAAGAGACGACACACAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1165  
 QY 185 -----ValSerPheCysSerSerSerSerSerSerSerSerSerSer 196  
 Db 1166 TTCACCAAGTAAACGGGTGGATCTTCTTAATCTTACCTCTTACCTGTGTTATGA 1225  
 QY 196 TGIUGLYLeuLys-----AsnTrpValAlaProThrArgGlnPheSerThrH 212

Db 1226 AATGGATCCGATGACGATGAACTTAACCGGCTCAGACCCCAATAACCACTTGCACCA 1285  
 QY 212 GGLGlnValLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 232  
 Db 1286 TCAAGTTCTTCCCTGAGATGATCTTAACGGCGG- 1327  
 QY 232 T-----LeuGlySerValGlyCysGlyGlnLeuGlnSerLe 244  
 Db 1328 TTCTGGCTTCTCTCGGCTCACTGGTTGGTGTAGTTGT- 1370  
 QY 244 uSerLeuSerMetSerProGlySerGlnSerSerCysValThrAlaProSerGlyThrAs 264  
 Db 1371 -----CACTCGAATCTAGCCACCGGATCGTCCGGGTAA 1405  
 QY 264 pSerValAlaValAlaSerAlaValValArgGlyHnAlaValLeuGlnGlnGlnProVa 284  
 Db 1406 ACCTACCAACAGTTGGCGCT-----GCCGATGTGAGCCGACACAGCCGTT 1450  
 QY 284 HnAsnArgLysSerLeuAspThrPheGlyGlnArgThrSerGlnTrpArgGlyValThrAr 304  
 Db 1451 G---AAAAGAGTCCGCGCTGACCAAGTCAGAGAGTCTAGATAGAGGTGTACGTT 1507  
 QY 304 gHnAsnArgTrpThrArgGlyArgGlyGlnAlaHnLeuTrp-AspAsn- 318  
 Db 1508 TTACCGGCGTACCGAGAGATGGAATCTCATATTGTGTAATATCTCATATTTTAATT 1567  
 QY 319 -----SerCysLysLeuGlnG 324  
 Db 1568 CGTAATGATGATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1627  
 QY 324 YGlnThrArgLysGlyArgGlnValTrpLeuGlyGlyTr- 337  
 Db 1628 TTCAGGAGACTG--CGAAGACAGATTACTAGTAATTAATTAATTAATTAATTAATTT 1685  
 QY 337 ----- 337  
 Db 1686 TTGTATTGTTGGTGTGAAGAAATGTCATATAATTATTAATTAATTAATTAATTAAG 1745  
 QY 337 ----- 337  
 Db 1746 TGAATTTCAGACTGTCATGACAGCTCGGATTTTCTCTCTTGACTCTCTATAT 1805  
 QY 338 -----AspMetGlnGlu-LysAlaAla- 344  
 Db 1806 TGAATGTTATTATTATTATTTTAAATAATACCGAAGAAATTTAATAAATTAATT 1865  
 QY 345 -----ArgAlaTrpAspLeuAlaAlaLeuLysTrpGlyProSer 358  
 Db 1866 AATTTGTTATTATTAATAGCATATGATAGAGCTGTATTAATTCGGAGAGTAGAA 1925  
 QY 359 ThrHnLeuAsnPheSerLeuGlnAsnTrpGlnValGlnLeuGlnGlu- 374  
 Db 1926 GCGGATCATATTTACATGAGATTAATGATGATGATGATGATGATGATGATGATGAT 1985  
 QY 374 ----- 374  
 Db 1986 TATAAATATATTGTTTATTAACGATTTTAAAGGTTGGAGATTAATTAATTAAT 2045  
 QY 375 -----MetLysAsnMetSerArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 390  
 Db 2046 TGAATTTATATGATACATTAATTAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2105  
 QY 391 SerSerGlyPheSerArgGlyAlaSerLeuTrpArgGlyValThrArgHnSerGlnHn 410  
 Db 2106 AGCAGAGCTTCCCTCGAG 2162  
 QY 411 GYArgTrpGlnAlaArgLLeuGlyArgValAlaGly- 422  
 Db 2163 GGTGTTGGAAGCTCGAATGGGTCAATTCTTAAGCAAAAGTAAATTTCTCTCATTTT 2222  
 QY 422 ----- 422



```

QY 338 -----AspMetGluGlu-LysAlaAla----- 344
Db 9916 TGAAGTTGTTATTTATTTATTTTAAATACCGGAAAGAAATTTATATTAATTAATTT 9857
QY 345 -----ArgAlaTyAspLeuAlaAlaLeuLysTyTrpGlyProSer 358
Db 9856 AATTTGTTTATTTATATAGAGATATGATAGAGCTGCTAATTAATTCGCGAGTAA 9797
QY 359 ThrHisIleAsnPhseSerIleGluValenTyrgValGlnLeuGluGlu----- 374
Db 9796 GCGGATATTCATTTCAACATTCGAAAGTATGATGATGATGATGATGATGATGATGAT 9737
QY 374 ----- 374
Db 9736 TATTAACATAATTTGTTTATTAACGATTTTTAAAGTTGGAGATTAAATATTAAT 9677
QY 375 -----MetLysAsnMetSerArgGlnGluTyValAlaHisIleuArgLys 390
Db 9676 TGAATTTTATAGATGCTAATTTTACCAAGAAAGCTTCGTAAGCTACTTCGCGACAA 9617
QY 391 SerSerGlyPheSerArgGlyAlaSerIleTyArgGlyValThrArgHisIleGlnHis 410
Db 9616 AGCACAAGGCTCCCTCCGAGAGATTGCAAGTATAGAGGTGTCACCTTG--CATAGTGT 9560
QY 411 GlyArgTrpGlnAlaArgIleGlyArgValAlaGly----- 422
Db 9559 GGTCTGTTGGAGACTCGAATGGCTCAATTCAGCCAAAGTAAATTTCTCATTTT 9500
QY 422 ----- 422
Db 9499 ATATGACTCGAAACCTCATTTTATGTTGTTATTTAACTTGAGTTTGTCTCTG 9440
QY 423 -----AsnLysAspLeuTyIleuGlyThrPheSerThrGlnGluAlaAla-- 438
Db 9439 AATCTTATAAATAGCATGCTTATTTGGCTTGTTCACACCGACGCTCGAAGCTGCTAG 9360
QY 438 ----- 438
Db 9379 GTAAATGCTTTTGTGTTGATTTCTACACACACATTTGTATTAATGTTTCTGCT 9320
QY 439 -----GluAlaTyAspValAlaAlaIleTy 447
Db 9319 ACTAATGATTTTCTATTTTATATATATACAGAGTTTCGATAAGCTGCAATCA 9260
QY 447 sPheArgGlyAlaAsnAlaValaThrAsnPheserArgIleTyAspValGluArgI 467
Db 9259 ATGTAACGGCAAGACGCGCTGACCACTTGTATTCGATTTTACATGAGGAA----- 9205
QY 467 eMetAlaSerSerAsnLeuAlaGlyGluLeuAla 479
Db 9204 -----CTCAAGCCCGTAAATTTGCT 9184

RESULT 8
US-09-313-294A-6727
; Sequence 6727, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laloudi, Raghnunach V.
; APPLICANT: Ico, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313, 294A
; NUMBER OF SEQ ID NOS: 1999-05-14
; SOFTWARE: PERL Program
; SEQ ID NO 6727
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature

```

```

; OTHER INFORMATION: Inbyte ID No. 6476212 700352302H1
; NAME/KEY: unsure
; LOCATION: 297
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-6727

Alignment Scores:
Pred. No.: 6.4e-19 Length: 301
Score: 263.00 Matches: 54
Percent Similarity: 77.38% Conservative: 11
Best Local Similarity: 64.28% Mismatches: 19
Query Match: 7.51% Indels: 2
DB: 4 Gaps: 0

US-10-024-632-2 (1-663) x US-09-313-294A-6727 (1-301)
QY 332 ValTyLeuGlyGlyTyAspMetGluGluLysAlaAlaArgAlaTyAspLeuAlaAla 351
Db 51 GCTTACTCGGGGCGCTATGACAGGAGAGGCAAGTCTGCTGCTATGACCTCGACGCT 110
QY 352 LeuLysTyTrpGlyProSerThrHisIleAsnPhseSerIleGluValenTyrgValGln 371
Db 111 CTCAAGTACTGGGCTCCGAGACTCTGCTCAACTTCCCTGCGAGGATTACTCCAGCGAG 170
QY 372 LeuGluGluMetLysAsnMetSerArgGlnGluTyValAlaHisIleuArgLysSer 391
Db 171 ATGCCGAGATGAGAGCGCGCTCCCGGAGAGTAACTGGCTCCCTCCCGCGAGAGC 230
QY 392 SerGlyPheSerArgGlyAlaSerIleTyArgGlyValThrArgHisIleGlnHisGly 411
Db 231 AGCGGCTTCTCCAGGGGCGCTCCCAAGTACAGAGGCGCTGCG--AGGCATCACCAAGCGG 289
QY 412 ArgTrpGlnAla 415
Db 290 AG-TGGANGCA 300

RESULT 9
US-09-198-119C-54
; Sequence 54, Application US/09198119C
; Patent No. 6417428
; GENERAL INFORMATION:
; APPLICANT: Thomasow, Michael
; APPLICANT: Stockinger, Eric
; APPLICANT: Jaglo-Ottosen, Kirsten
; APPLICANT: Gilmore, Sarah
; APPLICANT: Zhang, Daniel
; APPLICANT: Jiang, Cai-Zhong
; TITLE OF INVENTION: Plant Having Altered Environmental Stress Tolerance
; FILE REFERENCE: 19117.713 Seq List
; CURRENT APPLICATION NUMBER: US/09/198, 119C
; NUMBER OF SEQ ID NOS: 1998-11-23
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 54
; LENGTH: 898
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: bncbf5 gene

```

US-09-198-119C-54

Alignment Scores:

Pred. No.:	4.78e-07	length:	899
Score:	158.50	Matches:	62
Percent Similarity:	37.00%	Conservative:	33
Best Local Similarity:	22.71%	Mismatches:	97
Query Match:	4.53%	Indels:	7
DB:	4	Gaps:	7

US-10-024-632-2 (1-663) X US-09-198-119C-54 (1-898)

QY	355	ThHhIstlAanbheserilegiuAnuYrgiValaIgiNleuGiuglUmeIyAsmMet	378
DB	46	ACTAATCTACAGCTCAGCGCTTATCCAGTTTCTCAAAAAAGTATCTACAGTAACTACGTCTC	105
QY	379	Ser-----ArgingluYrgValaIahIleuYrgAglYsSer	392
DB	106	TCTACTTTTCTGAACTGCTCCGCTCCGAGACAGAGTCTCCGGTTAATACGAAGTGTGT	165
QY	393	GlyPheSerArgglYalaser	399
DB	166	GATTACACTTTTGGCGGAGCTGTCCCAAGAACTCTGTAAGAGAACTTTCAGAG	225
QY	400	-----IleYrgAglYalThArGhIshIsgInIleaglYArgTrpInAla	415
DB	226	ACAGCCCAACCCCATTTACAGAGAGTT-----GCTCCAGGAATCTCAGTAATGGTGTGT	282
QY	416	ArgIleaglYArGValaIagIyAnuYsAspLeuYrYleuGIYThrPheSerThGInglu	435
DB	283	GAACTGGAGGAGCAACCAAGAAATGTAGAAATTTGGCTTCGGAACCTTCAAAACAGCTGAG	342
QY	436	GIuIalalagIuIaYrgAspValaIahIleYsPheArgglYAlAsnAlValThr	455
DB	343	ATCCACACTGCTGTCTCAGACGTTGCGGCTTAACTCTCCGTAAGAGAGCGCTGCCTC	402
QY	456	AsnPhe-----AspIle	459
DB	403	AACTTCGCGCACTCGGCTTGGCGGCTCCGATCCCGAGACAGACTGGCCCAAGATATTC	462
QY	460	SerArg-----YrgAspValGIuArgIleMetAlSerSer	471
DB	463	CAGAAAGCTGTGCTGAGCCGCACTTGGCTTTTGGAGCCGGAAGAAAGTATACCAACAGC	522
QY	472	AsnIleuIaIagIyIuIeAlArGArgYlYsAspAspProArgAsnIyAsp	491
DB	523	AATGATCATGGCATGAACATGGCTTCTCAAGTTGAGGTTAAAGACACAGACGATCAATGAC	582
QY	492	IleAspYrAsnYsSerValValThrSerValAsnAsnIuIu-----	506
DB	583	CTGGACATCGAGAGACAGATATGTGAGGCTGTTTAAAGAGGAACAGAGAAAGGTTT	642
QY	507	-----ThValGIuIalGIalagIYAsnAsn	516
DB	643	TACATGGCGAGAGACAGACGGTTGTGGGTGTTTCCGAGGAACAGATGACAAAGG	702
QY	517	AsnGIuIaAspSerGIuTrPLysMet-----	525
DB	703	TTTTTACATGACAGAGAGTGGATGTTTCGGATCCGACCTTGTGGCTGATATGGCGCA	762
QY	526	--ValIeuPheAsnHleProSerGIuGIuGIuAlaAsnGIYAsnGIYSerAspGIu	544
DB	763	GGAGATGCTTACCGCGCTCCGCTCCGTAACATGGGGAACATATATGATGACTTCGAAGAGAT	822
QY	545	LysIleMetIaenYsGIYAsnYrYrgAsnSerAlPhe	557
DB	823	GCTGACATGAACCTCTGGAATTAATTAAGTACTCATATTTT	861

RESULT 10  
US-09-202-161B-6  
; Sequence 6, Application US/09202161B  
; Patent No. 6653533  
; GENERAL INFORMATION:

```

1  APPLICANT: Purdue Research Foundation
2  TITLE OF INVENTION: GENES ENHANCING DISEASE RESISTANCE IN PLANTS
3  FILE REFERENCE: 7024-371
4  CURRENT APPLICATION NUMBER: US/09/202,161B
5  PRIOR FILING DATE: 1999-06-14
6  PRIOR APPLICATION NUMBER: PCT/US97/10382
7  PRIOR FILING DATE: 1997-06-12
8  PRIOR APPLICATION NUMBER: 60/046,494
9  PRIOR FILING DATE: 1997-05-14
10 PRIOR APPLICATION NUMBER: 60/019,633
11 PRIOR FILING DATE: 1996-06-12
12 NUMBER OF SEQ ID NOS: 30
13 SOFTWARE: ASCII
14 SEQ ID NO: 6
15 LENGTH: 1405
16 TYPE: DNA
17 ORGANISM: Lycopersicon esculentum
18 IS-09-202-161B-6

```

Alignment Scores:

Pred. No.:	1.67e-06	Length:	1405
Score:	166.50	Matches:	47
Percent Similarity:	38.8%	Conservative:	16
Best Local Similarity:	29.01%	Mismatches:	54
Query Match:	4.47%	Indels:	45
DB:	4	Gaps:	4

US-10-024-632-2 (1-663) X US-09-202-161B-6 (1-1405)

CY		359	ThHtIILasPheserleGlusnynrGlnalGlnleugluMctLVAsmet	378
			:::::	388
Dd		510	ACGGAGATCAACTTATGCCCTGAACAACCATCGATCCGCACAGAAAACGAAGTCGTG	569
CY		379	SerArgIngluTVaLAlahisLeuaRgaRlyserSerGlyPheSerArgLYaLa	398
Dd		570	TCTCCGATTCTGACGTCACT-----CGTCGAAAAAG-----	602
CY		399	SerIleTyArgIGlVAlThrARHishtsiGlnhisGIyArgrTPGlnALargileGly	418
Dd		603	-----TTTAGAGCCTT-----CGTAAGAACCCTGGGCTCTTGCGCTCGAAGATTCCG	653
CY		419	ArgValAlagIYAshLyAsnLyAspleuTYrengIyThrpheSerThrgIngluGLuAla	438
Dd		654	GACCCGACCCGGGGAAACGGGTGTGGTGGTACTTATGACACCCACAGAAAGACGACT	713
CY		439	GIuALeTyAspyAlaLalaLaleyspheARgGlyALAsenALVaLIthRAsnPhasp	458
Dd		714	GTGCTTACGATAAACCTCGAGTAAAGCTCAAAAGCTCTGACGCCGTTACCAATTTCCG	773
CY		459	IleSerArgTyAspVALgluARgtIleMetAlaserSerAsnLeuLeuAlagIyluu	476
Dd		774	GTATCAACA-----	782
CY		479	ALAArgArgLYelYAspAsnspProhRgsnlysaPIleapTYrAsnLyseVal	498
Dd		783	-----ACGGCGGAGTA	793
CY		499	VAlthSerValAsnAnsgluThrValGlnValGlnAlagIYAasnAsnAsnGlu	518
Dd		795	ACGGTGACGGTTACGGAACCGAAACCGAGTCTGTGCCACGCGTGAGATPAAGCGAA	854
CY		519	AsnAsp	520
Dd		855	AACGAT	860

RESULT 11  
US-09-026-039-1  
; Sequence 1, Application US/09026039  
; Patent No. 6329567  
; GENERAL INFORMATION:  
; APPLICANT: Jofuku, K. Diane  
; APPLICANT: Okumuro, Jack K.  
; TITLE OF INVENTION: Methods for Improving Seed

NUMBER OF SEQUENCES: 103  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: IBM PC compatible  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA: US/09/026,039  
 APPLICATION NUMBER: US/09/026,039  
 FILING DATE: 19-FEB-1998  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/912,272  
 FILING DATE: 15-AUG-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/879,827  
 FILING DATE: 20-JUN-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/700,152  
 FILING DATE: 20-AUG-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bastian, Kevin L.  
 REGISTRATION NUMBER: 34,774  
 REFERENCE/DOCKET NUMBER: 023070-067230US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1669 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: -  
 LOCATION: 1..1669 /note= "canola APETALA2 (AP2) domain"  
 OTHER INFORMATION: containing (ADC) gene"  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: 1..532  
 OTHER INFORMATION: /note= "gene sequences 5' to start site"  
 US-09-026-039-1  
 Alignment Scores:  
 Pred. No.: 3.65e-06 Length: 1669  
 Score: 154.50 Matches: 86  
 Percent Similarity: 32.03% Conservative: 43  
 Best Local Similarity: 21.39% Mismatches: 95  
 Query Match: 4.41% Indels: 180  
 DB: 4 Gaps: 8  
 US-10-024-632-2 (1-663) x US-09-026-039-1 (1-1669)  
 QY 153 HlsmetserValGlnThrHisProTyrTyrSerGlyLeuAlaCysHlleglyLeuTyrGln 172  
 DB 551 CATATTTCTTGTCTTTGTTCTCTTTATTACAGAGAGATATGGAACCTTAACGAC 610  
 QY 173 AlaPro-----LeuGlnGlnGluThrThrLysGluThrHisValserAspCysSer 190  
 DB 611 TCACCTGATACACACCAAGATCCGACGCTAGATGAGAAACGGGCTGAGATGTACCAATC 670  
 QY 191 LeuMetProGlnMetThrGlnGlyLeuLysAsnTprValAlaProThrArgGluPheSer 210  
 DB 671 TCTATGATCATCATGACCAACGATGTCTGTCTCTCTCTCCCGGTACCCGGAATATTTT 730

QY 211 ThrHisGlnGlnValLeuGlnGlnGlnMetLysCysGlyMetGlyAsnGluArgAsnGly 230  
 DB 731 TCCG--AATCAATCATGAAACAG----- 752  
 QY 231 ValSerLeuGlySerValGlyCysGlyGlnLeuGlnSerLeuSerLeuSerMetSerPro 250  
 DB 753 -----GAACTTCAGAGATATGTGGTCCGATCACT 785  
 QY 251 GlySerGlnSerSerCysValThrAlaProSerGlyThrAspSerValAlaValAspAla 270  
 DB 786 AGAAACCGATCTCTTGTTCGGTCGATCTTCACCGGCTGTGT----- 827  
 QY 271 LysLysArgGlyHisAlaLysLeuGlyGlnLysGlnProValHisArgLysSerIleAsp 290  
 DB 828 ---CGTCCGAAACCTAGACCTGGAGATACACACGCCGTA---AAAAGACCGACGT 881  
 QY 291 ThrPheGlnArgThrSerGlnTyrArgGlyValThrArgHisArgTprThrGlyArg 310  
 DB 882 GGTCTCGCTACGAGCTCTCAGTATAGAGAGTTACTTTTATTCAGAACCGAAGA 941  
 QY 311 TyrGluAlaHisLeuTyr----- 316  
 DB 942 TGGAGTCAATTTTGGTAATTAATTCTTAACCGAGATATACCAATCACTATTA 1001  
 QY 316 ----- 316  
 DB 1002 TTACCTAATAGTAAATCATCAATATACATGTTTCATTTCATTGAGCCNAATACCGAT 1061  
 QY 317 -----AspAsnSerCysLysLysGlnGlnGlnThrArgLysGlyArgGlnValTyrLeu 334  
 DB 1062 GTTGTGTTTAAATATGTTTGAATCTTATGAGGAGCATGCG--GGAACCAAGTGTACTTA 1120  
 QY 335 Gly----- 335  
 DB 1121 GGTATGATCATGTAAATGTTGTTCAAACACAGATCAATATCTATTTGAAGTTGTG 1180  
 QY 335 ----- 335  
 DB 1181 TTGTGTCTGTCATTTTATATGATTTCTTCGACCAATTAAGTTTATATTCCTTA 1240  
 QY 335 ----- 335  
 DB 1241 TATTACTTTTGTACATATTCAGGTGATTTGACACACACATGCCGTGCTCGATAG 1300  
 QY 336 -----GlyTyrAspMetGlnGly----- 342  
 DB 1301 TTTTACTCATCCAAATATGATCAATTAAGCAATCTAATTCCTTATTTTGTAAATTG 1360  
 QY 343 -----AlaAlaArg-AlaTyrAspLeuAla 351  
 DB 1361 CTGATATACAAATTAATTTGGTGGGTAACGTTTGGGACAGTGCCTTACGATAGACCCC 1420  
 QY 351 AlaLeuTyrTyrPglProSerThrHisIleAsnPheSerIleGluAsnTyrGlnVal 371  
 DB 1421 AGTTAAGTTTAAAGGTGATGATGCAGATTAATTAATTAATTAAGTAAAGTATGTGAGA 1480  
 QY 371 nLeuGlnGlnMetLys----- 376  
 DB 1481 TTTGAACACGATAAATATTATTTATTTGTAGATTCAACAAATGACTTATTAATTACT 1540  
 QY 377 -----AsnMetSerArgGlnGluTyr 384  
 DB 1541 CGAATCAAAACAAATTAATTTTGTTCAGATGACGAGTTGACAAAGAAAGATTCA 1600  
 QY 384 lAlaHisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGly 404  
 DB 1601 GCATGTCATTAGAGCAAGCAAGACAGTGGTTCCAAAGGACAGCTTAAGATATAGAGTGT 1660  
 QY 404 lThr 405  
 DB 1661 CACT 1664  
 RESULT 12



```

Db      318 CCAACAAAAATCTAGATTGGCTGGAACTTCCAAACGCTGATGCGAGCTCGT 377
Qy      440 AlatyAspValAlaIleIleLysPheArgGlyAlaAsnAlaValThrAspHe----- 457
Db      378 GCTCACGACGCTGCGCTTCTAGCTCCGTGGAAGAGGCGCTGCTCACTTCCGCGAC 437
Qy      458 -----AspIleSerArg----- 461
Db      438 TCGGCTTGGCGCTCCGTATCCCGAGACAACTCGCCAGAGATATCCAGAGGCTGCT 497
Qy      462 -----TyrAspValGluArgIleMetAlaSerSerAsnLeuAla 475
Db      498 GCTGAAGCCGATTTGCTTTTGAAGCGCGAAGAGTATACACAGCAAGATGATCATGCGC 557
Qy      476 GlyGluLeuAlaArgArgLysAspAsnAspProArgAsnLysAspIleAspTyrAsn 495
Db      558 ATGAACATGGCTTCTAGAGCTGAGTTAATGACAGACGATCATGCTGACATGAG 617
Qy      496 LysSerValValThrSerValAsnAsnGluGlu----- 506
Db      618 GAGACGATGTGGAGGCTGTTTACTGAGAGCAGAGACGAGGTTTACATGCGCGAG 677
Qy      507 -----ThrValGluValGlnAlaGlyAsnAsnAsnGluAsnAsp 520
Db      678 GAGACGACGCTGAGAGGCTGTTTCCGAGAGAACAGATGACAAAGGTTTACATGAC 737
Qy      521 SerGluTyrLysMet-----ValLeuPhe 528
Db      738 GAGAGATGATGTTCGGATGCGGACCTTGTGGCTGATATGCGCGAGAGTCTTATA 797
Qy      529 AsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsnGlySerAspGluLysIleMetAsn 548
Db      798 CCGCCGCGCTCCCTACAAATGGGACATTAATGATGACTTGAAGAGATGCTGACATGAC 857
Qy      549 CysGlyAsnTyrArgAsnSerAlaPhe 557
Db      858 CTCTGGAATTATTAGTACTGATTTT 884

RESULT 14
US-09-198-119C-52
; Sequence 52, Application US/09198119C
; Patent No. 6417428
; GENERAL INFORMATION:
; APPLICANT: Thomashow, Michael
; APPLICANT: Stockinger, Eric
; APPLICANT: Jaglo-Octosen, Kirsten
; APPLICANT: Zarka, Daniel
; APPLICANT: Zhang, Cai-Zhong
; TITLE OF INVENTION: Plant Having Altered Environmental Stress Tolerance
; FILE REFERENCE: 19117,713 Seq List
; CURRENT APPLICATION NUMBER: US/09198,119C
; PRIOR APPLICATION NUMBER: US 08/706,270
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: US 09/018,233
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: US 09/017,816
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: US 09/018,235
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: US 09/017,575
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: US 09/018,227
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: US 09/018,234
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 874
; TYPE: DNA

```

```

; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: bnCBF4 gene
US-09-198-119C-52

Alignment Scores:
Pred. No.: 7,528-06
Score: 147.00
Percent Similarity: 37.80%
Best Local Similarity: 25.36%
Query Match: 4,204
Gaps: 5
Length: 874
Matches: 53
Conservative: 26
Mismatch: 78
Indels: 52

US-10-024-632-2 (1-663) x US-09-198-119C-52 (1-874)

Qy      400 ILETYRAAGLYValThrArgHisIleGlnIleGlyArgTyrGlnAlaArgIleGlyArg 419
Db      217 ATTATCCAGAGAGT---CGCTTAGAAAATCAAGTATAGGTGTGTGTAAGAGGAA 273
Qy      420 ValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheSerThrGlnGluAlaAlaGlu 439
Db      274 CCAACAAAAAATCTAGATTGGCTGGAACTTCCAAACGCTGATGCGAGCTCGT 333
Qy      440 AlatyAspValAlaIleIleLysPheArgGlyAlaAsnAlaValThrAspHe----- 457
Db      334 GCTCACGACGCTGCGCTTCTAGCTCCGTGGAAGAGGCGCTGCTCACTTCCGCGAC 393
Qy      458 -----AspIleSerArg----- 461
Db      394 TCGGCTTGGCGCTCCGTATCCCGAGACAACTGCGCAAGATATCCAGAGGCTGCT 453
Qy      462 -----TyrAspValGluArgIleMetAlaSerSerAsnLeuAla 475
Db      454 GCTGAAGCCGATTTGCTTTGAGCGCGAAGAGTATACACAGCAAGATGATCATGCGC 513
Qy      476 GlyGluLeuAlaArgArgLysAspAsnAspProArgAsnLysAspIleAspTyrAsn 495
Db      514 ATGAACATGGCTTCTAGAGCTGAGTTAATGACAGACGATCATGCTGACATGAG 573
Qy      496 LysSerValValThrSerValAsnAsnGluGlu----- 506
Db      574 GAGACGATGTGGAGGCTGTTTACTGAGAGCAGAGACGCGTTTACATGCGCGAG 633
Qy      507 -----ThrValGluValGlnAlaGlyAsnAsnAsnGluAsnAsp 520
Db      634 GAGACGACGCTGAGAGGCTGTTCCGAGAGAACAGATGACAAAGGTTTACATGAC 693
Qy      521 SerGluTyrLysMet-----ValLeuPhe 528
Db      694 GAGAGATGATGTTCGGATGCGGACCTTGTGGCTGATATGCGCGAGAGGATCTTATA 753
Qy      529 AsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsnGlySerAspGluLysIleMetAsn 548
Db      754 CCGCCGCGCTCCCTACAAATGGGACATTAATGATGACTTGAAGAGATGATGACATGAC 813
Qy      549 CysGlyAsnTyrArgAsnSerAlaPhe 557
Db      814 CTCTGGAATTATTAGTACTGATTTT 840

RESULT 15
US-09-533-029-21
; Sequence 21, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Brown, Pierre
; APPLICANT: Kiechmann, Jose-Luis
; APPLICANT: Keddie, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang

```



```

APPLICANT: Ratcliffe, Oliver
APPLICANT: Pilgrim, Marsha
APPLICANT: Jieng, Cal-Zhong
APPLICANT: Reuber, Lyne
TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
FILE REFERENCE: MBI-010
CURRENT APPLICATION NUMBER: US/09/533,029
EARLIER APPLICATION NUMBER: 60/125,814
EARLIER FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 121
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 1055
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G19
US-09-533-029-21

```

```

Alignment Scores:
Pred. NO.: 1.04e-05 Length: 1055
Score: 147.00 Matches: 67
Percent Similarity: 37.83% Conservative: 34
Best Local Similarity: 25.09% Mismatches: 100
Query Match: 4.20% Indels: 66
DB: 4 Gaps: 10

```

US-10-024-632-2 (1-663) x US-09-533-029-21 (1-1055)

```

QY 328 LysGlyArgGlnValTYrLeuGlyTYrAspMetGluGlnValaAlaArgAlaTYr 347
DB 121 AAGGCGGTAAATC-----ACGGCTGAGGAACTCTGTCAGAGCTCGAT 165
QY 348 AspleuAlaAlaLeuTYrTYrGly-----ProSerThr 359
DB 166 GCTTCGCGCGCGAGCACTCTGGGTTTCTATTCCACTCCAACTCCATCCACCAAC 225
QY 360 HisLeuAspPheSerIleGluAsnTYrGlnValGlnLeuGlnGluMetLysAsnMetSer 379
DB 226 CAAGTTAAC-----GTGAAGAGAGGACGACGTGAAGAAAGAGCAG 264
QY 380 ArgGlnGluTYrValAlaHisLeuArgGlySerSerGlyPheSerArgGlyAlaSer 399
DB 265 GCAACAGAGCGGGGAAACGAGGAAAGAGAG-----AAT 300
QY 400 IleTYrArgGlyValThrArgHisIleGlnHisGlyArgTYrGlnAlaArgIleGlyArg 419
DB 301 GTTATGAGGGATA---CGTAAGCGTCATGGGAAATGGCGGCTGAGATTTCGAGAT 357
QY 420 ValAlaGlnValAspLeuTYrLeuGlyThrPheSerThrGlnGlnGluAlaGln 435
DB 358 CCACGAAAGGTGTAAGGTTGGCTTGTGTAACAGGGCGAGGAAAGCTGCCATG 417
QY 440 AlaTYrAspValAlaAlaIleLysPheArgGlyValaAsnAlaValThrAspPheAspIle 459
DB 418 GCTTATGATGTTGGCGGACAGACATCGTGATTAAGCCAGCTCACTTCCAGAT 477
QY 460 SerArgTYrAspValGlnArgIleMetAlaSerSerAsnLeuAlaGlyGlnLeuAla 479
DB 478 CTGCACCATCTCT-----CCTCCTCTAATTATCTCTCCGCGTCATCG 522
QY 480 ArgArgGlyLysAspAsnAspProArgAsnLysAspIleAspTYrAsnLysSerValVal 499
DB 523 CCACGATCAACCGATCAGCTCCGCGCAAGAG-----GTC 558
QY 500 ThrSerValaAsnGlnGlnThrValGlnValaGlnAlaGlyAsnAsnAsnGlnAsn 519
DB 559 TGGCTTGTCT-----TCTCAGAGT 576
QY 520 AspSerGluTYrPlyMetValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGly 539
DB 577 GAGAGCGAG-----TTAAGTCAAGCGAGTTTCCCGGTGAGAGTGTATAGGA 621

```

```

QY 540 AsnGlySerAspGlnLysIleMetAsnCyseGly-----AsnTYrArgAsn 554
DB 622 TTGGAAATGGGACGAGGTTTCAAACTGAGTTACGGATTTGACCGGATTAATGATCTG 681
QY 555 SerAlaPheSerMetAlaLeuGlnAspLeuIleGlyTleAspSerValGlySerGlyGln 574
DB 682 AAACAGCAGATATCGAGCTTGGAATCGTCTTGAGCTGACGGTAAACAGCGCGACCAA 741
QY 575 HisAsnMetLeuAspGluSer 581
DB 742 CCGAGTCACTTGAATGAGTCC 762

```

Search completed: March 13, 2004, 08:01:55  
Job time : 131 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: March 13, 2004, 00:42:03 ; Search time 430 Seconds  
(without alignments)  
6550.129 Million cell updates/sec

Title: US-10-024-632-2  
Perfect score: 3502  
Sequence: 1 MKRINSENNITDDGNHNMWLG.....RSPALISLHPVFASTWDT 663

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-NODE=frame+.p2n.model -DEV=x1h  
-Q=/cgn2.1/USPFO spool/US-10024632/runat 09032004 101207 10754/app query.fasta\_1.839  
-DB=N Geneseq 23Jan04 -QFMT=fastaop -SUPFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -STRAP=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-NODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10024632 @CGN 1 1 352 @runat 09032004 101207 10754 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WALT -DSPELOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2002as:\*  
6: geneseqn2003as:\*  
7: geneseqn2003bs:\*  
8: geneseqn2003cs:\*  
9: geneseqn2003ds:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No	Score	Query Match	Length	ID	Description
1	3502	100.0	2344	6	ABN84480
2	1715	49.0	2323	6	ABN84481
3	1408.5	40.2	1926	6	ABN84482
4	1349.5	38.5	2010	6	ABN84484
5	1338	37.9	1668	3	AAC43005
6	1328	37.9	2148	3	AA559220
7	1328	37.9	2148	7	ABX13437
8	1307	37.3	1738	3	AA559222

9	1104.5	31.5	1954	7	ACC00850
10	1059.5	30.3	2463	7	ACC00803
11	1048	29.9	2588	7	ACC00801
12	1032	29.2	1290	7	ACC00805
13	977	27.9	7367	6	ABN84483
14	952.5	27.2	1874	7	ACC00799
15	951.5	27.2	1818	9	ADD30822
16	951.5	27.2	1818	9	ADD30822
17	951.5	27.2	1818	9	ADD30333
18	948	27.1	2011	4	ACC91399
19	945.5	27.0	1500	9	ADD30559
20	944	27.0	2014	4	ACC91398
21	938	26.8	1959	7	ACC00847
22	936	26.7	1990	7	ACC00806
23	935.5	26.7	1935	7	ACC00848
24	930	26.6	2052	9	ADD30980
25	927	26.5	1824	7	ACC00802
26	922.5	26.3	1864	7	ACC00791
27	922.5	26.3	1878	7	ACC00846
28	920.5	26.3	2168	7	ACC00804
29	920	26.3	1758	7	ACC00794
30	911.5	26.0	1753	7	ACC00792
31	894.5	25.5	2374	7	ACC00849
32	891.5	25.5	1941	9	ADD30349
33	718	20.5	551	7	ACC00796
34	701	20.0	1726	7	ACC00845
35	696.5	19.9	1271	7	ACC00797
36	684	19.5	1753	7	ACC00844
37	682	19.5	1465	6	ABQ81396
38	682	19.5	1530	9	ABK65225
39	682	19.5	1530	9	AD31800
40	682	19.5	1530	9	ADC46704
41	682	19.5	1530	9	AD30581
42	681.5	19.5	1539	6	ABQ81395
43	680	19.4	1619	7	ACC00843
44	678	19.4	1710	7	ACC00785
45	674	19.2	1065	9	ADD31036

## ALIGNMENTS

RESULT 1	ABN84480	standard; cDNA; 2344 bp.
ID	ABN84480	
AC	ABN84480;	
DT	21-OCT-2002	(first entry)
XX		
DE	soybean AINTEGUMENTA-like polypeptide GmANT1 cDNA.	
XX		
KM	AINTEGUMENTA; ANT-like polypeptide; GmANT1; soybean; transgenic plant;	
XX	plant; Gene; se.	
OS	Glycine max.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	242..223
FT		/*tag= a
FT		/product= "GmANT1"
XX		
XX	W0200259332-A2.	
XX		
PD	01-AUG-2002.	
XX		
PF	19-DEC-2001; 2001MO-US049294.	
XX		
PR	21-DEC-2000; 2000US-0257896P.	
XX		
PA	(MONS ) MONSANTO TECHNOLOGY LLC.	
XX		
PI	He SS, Dotson SB;	
XX		

DR MPI: 2002-599798/64.  
 DR P-PSDB, ABB79636.  
 XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in  
 PT improving agronomic, horticultural, and quality traits of plants, such as  
 PT increased size of plant organs.  
 XX  
 XX Example 1; Page 127-130; 169pp; English.

XX The present sequence is that of cDNA encoding GmANT1, a newly identified  
 CC AINTEGUMENTA-like (ANT-like) polypeptide of soybean. The cDNA comprises  
 CC the insert of clone CPR6763, which was obtained from a soybean sequence  
 CC database screening using Arabidopsis ANT as query. Also isolated was  
 CC clone CPR6762 (see ABBN4481) encoding GmANT2 (see ABB79637). GmANT1 and  
 CC GmANT2 show homology to ANT in 2 N-terminal Ap2 DNA binding domains, but  
 CC have C-terminal sequences that bear little, if any, homology to ANT. The  
 CC invention provides nucleic acids encoding ANT-like polypeptides  
 CC comprising, in the N-terminal to C-terminal direction, 2 Ap2 DNA binding  
 CC domains followed by an amino acid subsequence selected from those given  
 CC in ABB79623-35. ANT-like polypeptides have been identified in soybean,  
 CC rice, cotton and corn (see ABB79636-41). Nucleic acids (see ABBN4480-86)  
 CC encoding the ANT-like polypeptides can be used in the construction of  
 CC transgenic plants, especially corn, soybean, canola, wheat, cotton,  
 CC tomato or potato (all claimed) having improved agronomic, horticultural  
 CC or other quality traits, such as increased size of plant organs. These  
 CC plants are especially useful for production of ethanol or animal  
 CC feedstuff. The nucleic acids may also be used in antisense technology to  
 CC suppress endogenous ANT-like gene expression, in sense co-suppression to  
 CC modulate expression of endogenous ANT-like genes, as probes for  
 CC genetically and physically mapping the genes that they are part of, as  
 CC markers for traits linked to those genes, or in the identification of  
 CC loss of function mutant phenotypes of a plant due to a mutation in one or  
 CC more endogenous genes encoding the ANT-like polypeptides

XX Sequence 2344 BP; 738 A; 472 C; 548 G; 586 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3,966-284 Length: 2344  
 Score: 3502.00 Matches: 663  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-10-024-632-2 (1-663) X ABBN4480 (1-2344)

QY 1 MetLysArgLeuGlnSerAsnAsnThrAspArgLysAsnHisAsnTrpLeuGly 20  
 DB 242 ATGAAGGCAATGATGAGTACACACCGATGATGGAACACATCACTGTTGGGG 301  
 QY 21 PheSerLeuSerProHisMetLysMetGluAlaThrSerAlaAlaThrValProThrThr 40  
 DB 302 TTCTCTCTCCACCCCACTGAAATGAGAGCTACTTACGACCCCTGTTCCGACACC 361  
 QY 41 PheTyrMetSerProSerGlnSerHisLeuSerAsnPheGlyMetCysTyrGlyValGly 60  
 DB 362 TTCTACATGTCCTCTTCACTCACTGTCACATTCGAAATGCTTACGTTCCGA 421  
 QY 61 GluAsnGlyAsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeu 80  
 DB 422 GAAATGGTAACTTCCATCTTCCACTTACGGTTATGCTCTCAAGCTCATATGGGTCAC 481  
 QY 81 CysLeuLeuGluAlaLeuLysArgSerGlnThrGluValMetValProThrSerSerPro 100  
 DB 482 TGAATCTTGAAGCTTCAAAAGATCCAAACCGAAGTCAATGATGTCACCTTCGTC 541  
 QY 101 LysLeuGluAspPheLeuGlyAlaAlaThrMetGlyThrHisGluTyrGlySerHisGlu 120  
 DB 542 AAATTTGAGGACTTCTTACGTTGGTGCACATATGGGAACTCAACAAATATGGAAGCCAG 601  
 QY 121 ArgGlyLeuSerLeuAspSerTyrTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsn 140  
 DB 602 AGAGTTTGAGCTTGAACGATCATATTATTAATCCCAAAACGAGAGGCTCAACCAAC 661

QY 141 ArgAspLeuLeuSerGlnProPheArgGlnGlnGlnGlyHisMetSerValGlnThrHisPro 160  
 DB 662 AAGAGACCTTCTTCCAAACCCCTTACAGGCAACAGGCTCATATGATGCTCCAAACACACCT 721  
 QY 161 TyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGlnGlnThrThr 180  
 DB 722 TATTACTAGAGCCCTTGCTTGCATGCTTATATATCAAGACCGTTGGAGAAAGAAACA 781  
 QY 181 LysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlnGlyLeuLys 200  
 DB 782 AAGGAACCCACAGTGTGGATTCAGCTCCCTAATGCTTCAATATCAAGAACGCTTGAAA 841  
 QY 201 AsnTrpValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlnGlnMet 220  
 DB 842 AACTGGGTGCTCCAAACAGGAGATTTCACTCCAGACGAGTTTGGAGCAGCAATG 901  
 QY 221 AsnCysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGlyCysGlyGlu 240  
 DB 902 AATTGGCATGGGAGATGAGAAATGATGTGTCTTAAAGATCTGTGGGTGTGAGAG 961  
 QY 241 LeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValThrAlaPro 260  
 DB 962 TTCAATCTCTAAGCTTATCTTATGACTCTGTCTTCACTAAGTTGTGTCTGCTCT 1021  
 QY 261 SerGlyThrAspSerValAlaValAspAlaLysLysArgGlyHisAlaLysLeuGln 280  
 DB 1022 TCTGAAACAGATTTCTTCTGCTGTGATGCAAGAAAGAGGCGCATGCTAACTGTCTCG 1081  
 QY 281 LysGlnProValHisArgLysSerLeuAspTrpPheGlyGlyArgThrSerGlnTyrArg 300  
 DB 1082 AACACAGCTGTGATCATGAAATTTATTCGACACATTTGGGCAAGAACCTCGCAGTATAC 1141  
 QY 301 GlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHisLeuTrpAsnSerCys 320  
 DB 1142 GGTGTCAACAGCATATGATGAGATGAGTGTAGTATGAAGCCCATTTGGGATATATCTTC 1201  
 QY 321 LysLeuGlnGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlu 340  
 DB 1202 AAGAGGAAGGGAACCTGAGAAAGGACGACAGATATTTGGGGGTTATGATATCGAG 1261  
 QY 341 GluLysAlaAlaArgLysTyrAspLeuAlaAlaLeuLysTyrTrpGlyProSerThrHis 360  
 DB 1262 GAGAAAGCTGCAAGAGCTTATGATCTCGGGCCCTTAAGTACGCGGACCTTCAACGCA 1321  
 QY 361 IleAsnPheSerIleGluAsnTyrGlnValGlnLeuGlnGlyMetLysAsnMetSerArg 380  
 DB 1322 ATAAACCTTTCGATAGAAATTCACAAAGTTCACTTGAGGAATGAAACATGACGACA 1381  
 QY 381 GlnGluTyrValAlaHisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIle 400  
 DB 1382 CAGGAATACCTTTCACACTTGAGAAAGAAAGACAGCGGGTTTCTAGAGGTGCTTCATAT 1441  
 QY 401 TyrArgGlyValThrArgHisHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgVal 420  
 DB 1442 TACAGAGGGGTTCACAAAGCATCACCAAGAGAGTGGCAACGAGGATAGGAGAGTT 1501  
 QY 421 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheSerThrGlnGlnGluAlaAlaGluAla 440  
 DB 1502 GCTGGGAACAAAGACCTTTCCTTGGGAGCTTCAGCACCCCAAGGAGGAGCAGCAAGCA 1561  
 QY 441 TyrAspValAlaAlaAlaLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSer 460  
 DB 1562 TACATGTAGGGGAGTCAAAATTTCCGGGGCAATGCAATCAACAACTTGTGACATTTCA 1621  
 QY 461 ArgTyrAspValGluArgIleMetAlaSerSerAsnLeuAlaGlyGluLeuAlaArg 480  
 DB 1622 AGATACGATGTGGAGAAATCAATGCGCAATAGCAATCTCTCGCTGGGAGCTTGCAAG 1681  
 QY 481 ArgLysLysAspAsnAspProArgAsnLysAspIleAspTyrAsnLysSerValValThr 500  
 DB 1682 CGTAAGAAAGATTAACGATCTCTAGAAACAGAGCATAGACTTACCAACAGAGTGTAGTACA 1741

```

QY 501 SerValAsnGluGluThrValGlnValGlnGlyAsnAspAsnGluAsp 520
DB 1742 AGTGTGAACAATGAGAAAGGTTCAAGTCAAGAGAAACAATATGAAACGAC 1801
QY 521 SerGluTrpIleMetValLeuPheAsnIleProSerGlnGlnGlnIleAsnGlyAsn 540
DB 1802 TCAGAGTGAAGATGTTTATTTAACCACTTCACAGACCAAGCAATGGCAAT 1861
QY 541 GlySerAspGlnIleMetAsnGlyAsnTrpIleAsnSerIlePheSerMetAla 560
DB 1862 GCGAGTACCAAAAATAATGAACTGTGAAATTAACAGAAACAGTGCATTTCTATGAGCC 1921
QY 561 LeuGlnAspLeuIleGlyIleAspSerValGlySerGlyGlnIleAsnMetLeuAspGlu 580
DB 1922 CTACAAAGATCTTATGAGATTTGATTCGTGGGTTCTGGGACATATATGCTGACAG 1981
QY 581 SerSerIleGlyIleThrIlePheSerAsnTrpSerSerLeuValThrSerLeuSerSer 600
DB 1982 TCTAGCAAAATGGGACTCATTTTTCAAACACGTCATCGCTGGTGAACAAGTTTAAACACG 2041
QY 601 SerTrpGluIleAspProGluIleValArgGlyProSerLeuLeuPheProMetProProMet 620
DB 2042 TCAAGAGAGGCTAGTCTGAGAAAAAGGGTCCCTGCTGTGTTCCCAATGCTTCCAAAG 2101
QY 621 GluThrIleValIleAsnProIleGlyThrSerValThrSerTrpLeuProSerProThr 640
DB 2102 GAAACAAAGATTGTGAACCCCATTTGTAACAGTGTTCCTTGGCTACCTCACCAACG 2161
QY 641 ValGlnMetArgProSerProAlaIleSerLeuSerIleLeuProValPheAlaSerTrp 660
DB 2162 GTTCAAAATGAGGCTTCTCCGCTATCTCTTGTCTCACTTGCACATTTTGTCTTGG 2221
QY 661 ThrAspThr 663
DB 2222 ACTGACT 2230

```

RESULT 2  
ID ABN84481 standard, cDNA; 2232 BP.  
XX ABN84481;  
AC 21-OCT-2002 (first entry)  
DT 21-OCT-2002 (first entry)  
XX  
DE Soybean AINTEGUMENTA-like polypeptide GmANT2 cDNA.  
XX  
KM AINTEGUMENTA: ANT-like polypeptide; GmANT2; soybean; transgenic plant;  
KM plant; gene; ss.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 25..2022  
FT /tag= a  
FT /product= "GmANT2"  
XX  
XX WO200259332-A2.  
XX  
XX 01-AUG-2002.  
XX  
XX 19-DEC-2001; 2001WO-US049294.  
XX  
XX 21-DEC-2000; 2000US-0257896P.  
XX  
XX (MONS ) MONSANTO TECHNOLOGY LLC.  
XX  
XX He SS, Dotson SB;  
XX  
XX WPI: 2002-599798/64.  
XX P-FSDB; ABB79637.  
XX  
XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in  
PT improving agronomic, horticultural, and quality traits of plants, such as

```

PT increased size of plant organs.
XX
XX Example 1; Page 134-137; 16pp; English.
XX
XX The present sequence is that of cDNA encoding GmANT2, a newly identified
CC AINTEGUMENTA-like (ANT-like) polypeptide of soybean. The cDNA comprises
CC the insert of clone CP87626, which was obtained from a soybean sequence
CC database screening using Arabidopsis ANT as query. Also isolated was
CC clone CP87663 (see ABB79636). GmANT1 and
CC GmANT2 show homology to ANT in 2 N-terminal AP2 DNA binding domains, but
CC have C-terminal sequences that bear little, if any, homology to ANT. The
CC invention provides nucleic acids encoding ANT-like polypeptides
CC comprising, in the N-terminal to C-terminal direction, 2 AP2 DNA binding
CC domains followed by an amino acid subsequence selected from those given
CC in ABB79629-35. ANT-like polypeptides have been identified in soybean,
CC rice, cotton and corn (see ABB79636-41). Nucleic acids (see ABB84480-86)
CC encoding the ANT-like polypeptides can be used in the construction of
CC transgenic plants, especially corn, soybean, canola, wheat, cotton,
CC tomato or potato (all claimed) having improved agronomic, horticultural
CC or other quality traits, such as increased size of plant organs. These
CC plants are especially useful for production of ethanol or animal
CC feedstuff. The nucleic acids may also be used in antisense technology to
CC suppress endogenous ANT-like gene expression, in sense co-suppression to
CC modulate expression of endogenous ANT-like genes, as probes for
CC genetically and physically mapping the genes that they are part of, as
CC markers for traits linked to those genes, or in the identification of
CC loss of function mutant phenotypes of a plant due to a mutation in one or
CC more endogenous genes encoding the ANT-like polypeptides
XX
XX SO Sequence 2232 BP; 746 A; 482 C; 492 G; 603 T; 0 U; 0 Other;

```

Alignment Scores:

	Pred. No.:	Length:	Score:
1	5,51e-134	2323	1715.00
25	1715.00	378	65.86%
Best Local Similarity:	54.00%	Mismatches:	83
Query Match:	48.97%	Indels:	74
DB:	6	Gaps:	20

US-10-024-632-2 (1-663) X ABN84481 (1-2323)

```

QY 1 MetIySArgIleAsnGluSerAsnAsnThrAspAspGlyAsnAsnHisAsnTrpLeuGly 20
DB 25 ATGAGAGATGAGAAATGATGACATGCTGACCTTAATCAAAACAATGGTGGT 84
QY 21 PheSerLeuSerProHisMet-----LysMetGluAlaThr 32
DB 85 TTCTCACTCTCTCCCAATGCAATATATAGAGTTTCTTCACTCACACCTTCCCT 144
QY 33 SerAlaAlaThrValProThrThrPheTrpMetSerProSerGlnSerHisLeuSerAsn 52
DB 145 GCGCTGAGAGTGGTCTCTACAGCTTTTACCACACACGCTCCA-----CTTAGTAGC 198
QY 53 PheGlyMetCysTrpGlyVal---GlyGluAsnGlyAsnPheHisSerProLeuThrVal 71
DB 199 TATGTTTCTACTATGAGCTTGAAGCTGAATGTTGATTGTATTCAGCTTTGCCAATC 258
QY 72 MetProLeuIleSerAspGlySerLeuCysIleLeuGluAlaLeuLysArgSerGlnThr 91
DB 259 ATGCCCTCAATCTGATGCTCTCTCTATGATTTGAAACTTTAAGCAGGTCACAAAGCA 318
QY 92 GlnValMetValProThrSerSerProLysLeuGluAspPheLeuGlyIleAlaThrMet 111
DB 319 CAGCAATGGCTACTCTCAACACCAAACTGGAGAACTTCTTAGTGGGGAAGCCATG 378
QY 112 GlyThr---HisGluIle-----GlySerHisGluArgGlyLeuSerLeuAspSer 127
DB 379 GGGACCCCTCATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 438
QY 128 IleTrpTrpAsnSerGlnAsnAlaGluAlaGlnProAsnArgAspLeuLeuSerGlnPro 147
DB 439 GTTTTTCAC---ATCCAAACCTCACCGCGTGAACCCCAAAATTAATACCAAACTTACCAAAAC 495

```

QY 148 PheArgGln-----GlnGlyHisMetSerValGlnThrHisProTyrTyr 162  
 Db 496 CATGTTCAACACATTAGACCAACCAACCAACAGAGGCTTCAAGCATATAC 555  
 QY 163 SerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGlnGlnTyrThrLysGlu 182  
 Db 556 TCTACCTTGAGAAACCATGATATATATTA-----GAAAGGCTCAAGCA 600  
 QY 183 ThrHisValSerAspCysSerSerLeuMetProGln----- 194  
 Db 601 AGCCAACTCTGCAACACAACTTCATGTTCAAAAACATGGGTGATGATGCCGTT 660  
 QY 195 MetThrGlnGlyLeuLeuAsnTyrValAlaProThrArgGlnPheSer-----Thr 211  
 Db 661 CCTGTTCTGGCTTCAGAGTTGG-----GAAGTGAGAACTTCCAAAGCTAGCCATGCA 714  
 QY 212 HisGlnGln-----ValLeuGlnGlnMetLeuPheCysGlyMetGlyAsnGluArg 228  
 Db 715 CATGAGTCAAAGATGATGTTCTTCATGTGAGAGAAATGCTGCTGATCAGG- 768  
 QY 229 AsnGlyValSerLeuGlySerValGlyCysGlyLeuGlnSerLeuSerMet 248  
 Db 769 -----TCCATGATGATCATGAGCTTATGCTGATGCTGCAATGCTGATGCTTCATG 819  
 QY 249 SerProGlySerGlnSerSerCysValThr-----AlaProSerGlyThr 263  
 Db 820 AGCTCTAGCTCTCACTAGCAAGTGTGCAAGTCTCAGCGTCTCAGCTGCTGCTGCT 879  
 QY 264 AspSerValAlaValAlaAspAlaLysArgGlyHisAlaLysLeuGlnLysGlnPro 283  
 Db 880 GATTCCTGTTGCTCATGATGATCTAAGAAAGGGGCTGAAAGGTTGACCAACACCAAT 935  
 QY 284 ValHisArgLysSerHisAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValThr 303  
 Db 940 GTTCAATGAGAGTCCATGACCTTGTGCAAAAGAACCTCCAGTATAGAGAGTACA 999  
 QY 304 ArgHisArgTyrThrGlyArgTyrGlnAlaHisLeuTyrAspAsnSerCysLysGln 323  
 Db 1000 AGGCATAGAGTGGAGTGGAGATGATGAGCTCATCTTGGGACACAGCTGCAAGAAAGAG 1059  
 QY 324 GlyGlnThrArgLysGlyValArgGlnValTyrLeuGlyGlyTyrAspMetGlnLysAla 343  
 Db 1060 GGACCAAGCAGAAAGAAAGAAAGCAAGTTATCTAGGGGGTTATGATAGAAAGAAAGCT 1119  
 QY 344 AlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrProSerThrHisIleAsnPhe 363  
 Db 1120 GCGAGAGCTTATGATCTAGCGGACCTCAAGTATTGGGGACCTCCACTCACATATACCTT 1179  
 QY 364 SerIleGlnAsnTyrGlnValGlnLeuGlnMetLysAsnMetSerArgGlnLysTyr 383  
 Db 1180 CTTTGGAATAATATATATAATGAACTTGAAGAAATGAAACATGACCTAGCAAGATAT 1239  
 QY 384 ValAlaHisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGly 403  
 Db 1240 GTTGCTCATTTGAGAGAAAGAAAGACGCGATTCTCAAGAGGGGCTTCATGATACAGAGA 1299  
 QY 404 ValThrArgHisHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsn 423  
 Db 1300 GTAAACAAGACACACACATGAGAGTGGCAAGCTCGAATGTGAGAGTGGCTGGAAC 1359  
 QY 424 LysAspLeuTyrLeuGlyThrPheSerThrGlnGlnAlaGlnAlaGlyLysAspVal 443  
 Db 1360 AAAGATCTATATCTTGAACCTTTATACAAAGAGAGAGAGCTGAGCTATGATATT 1419  
 QY 444 AlaAlaIleLysPheArgGlyValAlaAsnAlaValThrAspPheAspIleSerArgTyrAsp 463  
 Db 1420 GCTGCTATATAATTCAGAGAGCGAATGCTGTACCAACTTTCACATCCACAGATATGAT 1479  
 QY 464 ValGlnArgIleMetAlaSerSerAsnLeuAlaGlyLysLeuAlaArgArgLysLys 483  
 Db 1480 GTGAGAAATCATGACAGACCAACCTCTTACGAGTGAAGTACTGAGGCAACGCA 1539  
 QY 484 AspAsnAspProArgAsnLysAspIleAspTyrAsnLysSerValValThrSerVal--- 502

Db 1540 GAGAGCGACATGAAACCTCAGTGCATTCATCAAAATACATCAATTAAGCTTCTGCATATGAG 1599  
 QY 503 AsnAsnGlnGlnThrValGlnValGlnAlaGlyAsnAsnAsnArgLysAspSerGlu 522  
 Db 1600 GACACTCAAGAGACCTATCTTAATGCAAGAGAGCTGTAGAGCGAAATGAT---CAG 1656  
 QY 523 TrpLysMetValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySer 542  
 Db 1657 TCGAAGATGTTCTCTAC---CAATCCTCAGCAACTGAGAGATCCACCAACATTT 1713  
 QY 543 AspGlnLysIleMetAsnCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGln 562  
 Db 1714 GAG-----AGTACAGAACTAACCAAGTCTCTTCGACGTGGCTTGGAC 1755  
 QY 563 AspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerSer 582  
 Db 1756 AACATGTT-----CATCAGAGAGTAGAGAAATCAAGT 1788  
 QY 583 LysIleGlyThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerArg 602  
 Db 1789 AAGCGAGAGACCATGTGTCAAAATCCTTCATTTGCGCACAAAGTTTGAGCAGCTCAAGA 1848  
 QY 603 GluAlaSerProGlnLysArgGlyProSerLeuLeuPheProMetProProMetGluThr 622  
 Db 1849 GAAGGTAGCCCTGATAGCAAGCTTCCAAATGCTCTCTGGAATGCCCTTCAACTGCATCA 1908  
 QY 623 LysIleValAsnProIleGlyThrSerValThrSerTyrPleuProSerProThrValGln 642  
 Db 1909 AAACATTTGGCTACTATATCAATTAACGTGAATTTCTTGGACCTTCAACCCCATTTGAGG 1968  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTyrThrAsp 662  
 Db 1969 -----CCAGCACTTACTTTGCTCCTCAAAATGCCAGTTTTCGACCTTGACAGAT 2016  
 RESULT 3  
 ID ABN84482 standard; cDNA; 1926 BP.  
 XX AC ABN84482;  
 XX DT 21-OCT-2002 (first entry)  
 XX DE Rice AINTEGUMENTA-like polypeptide OsANTI cDNA.  
 XX KW AINTEGUMENTA; ANT-like polypeptide; OsANTI; rice; transgenic plant;  
 XX KM plant; gene; ss.  
 XX OS Oryza sativa.  
 XX OS Location/Qualifiers  
 XX FT Key 1..1926  
 XX FT CDS  
 XX FT /feature="a  
 XX FT /partial  
 XX FT /product="OsANTI"  
 XX FT /note="the CDS does not include a stop codon"  
 XX PR 21-DEC-2000; 2000US-0257896P.  
 XX PA (MONS ) MONSANTO TECHNOLOGY LLC.  
 XX PI He SS, Dotsen SB;  
 XX WP1; 2002-595798/64.  
 XX DR P-PSDB; ABB79638.  
 XX XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in







QY 87 s-----ArgSerGlnThrGlnValMetValProThSerSerProLysLeuGluAspH 105  
 Db 345 GGGAGCGCAACCAAGCAAGAGGAGTGTGTGCGGCTCCGCCCAACCTGGAGATTT 404  
 QY 105 eLeuGly---GlyAlaThrMetGlyThrHisGlyTyrGlySerHisGlyValGlyLeuSe 124  
 Db 405 CCAAGGCGGCGCCCGGCGATG-----GCCCTGAG 434  
 QY 124 tLeuAsp-----SerLeuTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsnArgAs 142  
 Db 435 CCAAGCAACCTCGGCTTCTACTAGAGCGGCGCCGACCGTACACC----- 477  
 QY 142 tLeuLeuSerGlnProPheArgGlnGlnGlnMet-SerValGlnThrHisProTyrT 162  
 Db 478 -----CAAGGACACGCCCAAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 527  
 QY 162 ySerGlyLeuAlaCysHisGlyLeuTyrGln---AlaProLeuGlnGlnGlnGlnGln 181  
 Db 528 CGGCGCGCGCGCTTCTGCAAGTGTGCGTATCCCGCGCGCGCGCGCGCGCGCGCGCGCG 587  
 QY 181 ySerGlnThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlnGlyLeuLysA 201  
 Db 588 GCTG-GTGACCAACCACTCCCGCGCGCGAGT-----GCCGCCG 625  
 QY 201 snTPValAlaProThrArgGlnPheSerThrHisGlnGlnValLeuGlnGlnMetA 221  
 Db 626 GCTGGCGCGCG-----ATGCAC----- 642  
 QY 221 snCysGlyMetGlyLeuGlnArgGlnGlyValSerLeuGlySerValGlyCys----- 238  
 Db 643 -----GGCGCGCGCTTACCAATCGCAACGCCCGCGCGCGCGCGCGCGCGCGCGCG 694  
 QY 239 -----GlyLeuLeuGlnSerLeuSerLeuSerMet---SerProG 251  
 Db 695 GCCCATCATCCCAACCGCGCGCGCACTGCACCTCTCACCTCTGCATGACGTACGCGCG 754  
 QY 251 ySerGlnSerSerCysValThr-----AlaProSerGlyThrAspSerValAlaV 268  
 Db 755 GGTCCCACTCCAGCTGCGTCACTGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 814  
 QY 268 alaAspAla-----LysLysArgGlyHisAla---LysLeuGlyGlnLysGlnProValH 285  
 Db 815 TCGACCGCGCGAGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 874  
 QY 285 tAsnGlySerTlleAspThrPheGlyGlnAlaGlnThrSerGlnTyrArgGlyValAlaThrArgH 305  
 Db 875 ACCGCAAGTCCATGTACAGCTTCCGCGCGCGCAAGAGCGCGCGCGCGCGCGCGCGCGCG 934  
 QY 305 tAsnGlnThrPheGlyArgTyrGlnAlaHisLeuTyrAspAsnSerCysLysLysGlnGlyG 325  
 Db 935 ATAGGTGACTGGAGATATGAGGACACTCTGGGACAACTCACTCAAGAGAGAGGCGC 994  
 QY 325 tThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnGlnLysAlaAlaA 345  
 Db 995 AGACCGAAGAGAGCGCGCAAGGTATCTTGTGGGTATGACATGAGAGAGAGAGCGTCCA 1054  
 QY 345 tGAlaTyrAspLeuAlaAlaLeuLysTyrTPGlyProSerThrHisIleAsnPheSerI 365  
 Db 1055 GGGCGATATGATCTTGTGCGCTCAAGTACTGGGCGCTTCACGCACTCACTTCGCT 1114  
 QY 365 tLeuLeuTyrGlnValGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG 1174  
 Db 1115 TCGAGGACTCCAGAGAGAGCTGGAGAGATGAAGAACTAGAGAGAGAGAGAGTGTGG 1174  
 QY 385 tAsnLeuArgArgLysSerSerGlyPheSerArgGlyValAsnSerTlleTyrArgGlyValT 405  
 Db 1175 CTGACCTCAAG 1234  
 QY 405 tThrArgHisGlnGlnHisGlyTyrGlnAlaAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 425  
 Db 1294

Db 1295 ACCTTTCCTGGGACATTCATCCGCGCGCTTTTCCCGCGCGCGCGCGCGCGCGCATG 1354  
 QY 432 --SerThrGlnGlnGlnAlaAlaGlnAlaTyrAspValAlaAlaIleLysPheArgGlyA 451  
 Db 1355 CCGGCAAGCAG 1414  
 QY 451 tAsnAlaValAlaThrAsnPheAspIleSerArgTyrAspValGlnArgIleMetAlaSerS 471  
 Db 1415 TCAACGCGGTCAACCACTTCGACATCAGAGAGTACAGCTGGACAGAGATCTCTGGAGAGCA 1474  
 QY 471 eAsnLeuLeuAlaGlyGlnLeuAlaArgArgLys---LysAspAsnAspProArgAsnL 490  
 Db 1475 GCAGCTCTCTCCGGGAGAGTGGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1533  
 QY 490 yAspIleAspTyrAsnLysSerValAlaThrSerValAlaAsnGlnGlnThrValGlnL 510  
 Db 1534 -----GCCGCGCGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1567  
 QY 510 aGlnAlaGlyAsnAsnAsnAsnGlnAsnAspSerGlnTyrLysMetValLeu----- 527  
 Db 1568 TCGACGCGCGG-----AACGTGGCGAGTGGAGAGTGGCCACCGCGCGCGCG 1612  
 QY 528 -----PheAsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsnGlySerAspGln 545  
 Db 1613 CGCTGCCAGCGCGCGCGCGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1671  
 QY 545 yIleMetAsnCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuI 565  
 Db 1672 -----CACCATGACCTCTGCTCCGAGCGAGCGCTTCTGCTG---CTGCAGAGACATCG 1720  
 QY 565 tGlyLys---IleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlnSerSerLysI 584  
 Db 1721 TGTGACCGGTGAGACG 1755  
 QY 584 tGlyThrHisPheSerAsnThrSerSerLeuValAlaThrSerLeuSerSerPheArgGlnA 604  
 Db 1756 --GGCGCGCACTGTCATGCGCGCG-----ACAGGCTGGGCACTCTCCGGAGAC 1804  
 QY 604 tAspProGlnLysArgLysPro-----SerLeuLeuP 615  
 Db 1805 AGAGCGCTGAC---AGGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1861  
 QY 615 heProMetProProMetGlnThrLysIleValAsnProIleGlyThrSerValThrSerT 635  
 Db 1862 TCGCCAGCGCGCGCGCGCGCGCGCGCGCGAGCTGACAGCGCGGTG-----CCGCTGAACACT 1915  
 QY 635 tPLeu---ProSerProThrValGlnMetArgProSerProAla---IleSerLeuSerH 653  
 Db 1916 GGGCGTCCCTTCCCGCGCGGTAGCTCGGTGCGAGGCGCGCGCGCGCGCGCGCGCGCG 1975  
 QY 653 tLeuProValPheAlaSerTyrThrAsp 662  
 Db 1976 AACTGCCAATGTTGCCCGGTGACCGAC 2004  
 RESULT 5  
 AAC43005  
 ID AAC43005 standard, DNA, 1668 BP.  
 AC AAC43005;  
 DT 17-OCT-2000 (first entry)  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 37674.  
 XX Hybridation assay; genetic mapping; gene expression control;  
 XX protein identification; signal transduction pathway; metabolic pathway;  
 KM promoter; termination sequence; ss.  
 XX Arabidopsis thaliana.  
 OS  
 PN EP1033405-A2.

Wed Mar 17 08:18:55 2004

us-10-024-632-2.rng

Page 8

PD 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-013548P.  
PR 23-MAR-1999; 99US-0135786P.  
PR 25-MAR-1999; 99US-016264P.  
PR 29-MAR-1999; 99US-0162785P.  
PR 01-APR-1999; 99US-0177452P.  
PR 06-APR-1999; 99US-018224P.  
PR 08-APR-1999; 99US-018274P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132455P.  
PR 06-MAY-1999; 99US-0132466P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135114P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137503P.  
PR 07-JUN-1999; 99US-0137722P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139493P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139889P.  
PR 23-JUN-1999; 99US-0140383P.  
PR 23-JUN-1999; 99US-0140384P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140821P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141844P.  
PR 01-JUL-1999; 99US-0142154P.

PR 06-JUL-1999; 99US-0142290P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 27-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145921P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148665P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-015263P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.



[illegible][illegible]

QY 142 AspleuenserGlnProphe-----ArgGlnGlnGlnYHismetSerValGln 157  
 Db 719 ACTACAACTTTCAAGAGTTCTTACGTTCCCTCAACCGAAGAACAT----- 766  
 QY 158 ThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGln 177  
 Db 767 -----GAGGAA 772  
 QY 178 GluThrThrLysGluThrHisValSerProCysSerSerLeuMetProGlnMetThrGln 197  
 Db 773 GAAACT-----AGAAATTACGGGAATAC-----CCTAGTTTGACACAT 811  
 QY 198 GlyLeuLysAsnTrpValAlaProThrArgLupPheSerThrHisGlnGlnValLeuGln 217  
 Db 812 GGA----- 814  
 QY 218 GlnGlnMetAsnCysGlyMetClyAsnGlnArgGlnGlyValSerLeuGlySerValGly 237  
 Db 815 GGGCTTTTAAATGATGGGGATATATGGGGAATTTCA----- 850  
 QY 238 CysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysVal 257  
 Db 851 -----CAGTCACTGAGCTTATTCATGAGCCCTGGGTCACATTCAGTGCATC 898  
 QY 258 ThrAlaProSer----- 261  
 Db 899 ACTGGCTCACACACCAACCAAAACCAAAACCAACCAAGCCAAACACACAG 958  
 QY 262 -----GlyThrAspSerValAlaVal 268  
 Db 959 CAGATCTCGAAGCTCTGTGTGAGACAAAGCCTGGGTTTGAGACGACACATGGCGCT 1018  
 QY 269 AspAlaLysLysArgGlyHis-----AlaLysLeuGlyGlnLysGlnProValHis 285  
 Db 1019 GGGAGAAAGAAAGGGGACAAAGAGATGTGATGTTGGTCAGAAACAGATTGTCAT 1078  
 QY 286 ArgLysSerIleAspThrThrGlyGlnArgThrSerGlnTyrArgGlyValThrArgHis 305  
 Db 1079 AGAAATCTATCGAATACCTTTTGACAAAGCACTTCATACCGAGGGGTTCACAGACAT 1138  
 QY 306 ArgTrpThrGlyArgTyrGlnAlaHisLeuTrpAspAsnSerCysLysLysGlyGln 325  
 Db 1139 AGATGACCTGATGATATGATGATCTCATATGCGACATATGTTCAAGAGAGATGCAC 1198  
 QY 326 ThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnGlnLysAlaArg 345  
 Db 1199 AGTAAAGAAAGAAAGCAAGATTTATCTGGAGGTTTGATGATGAGAGAGAAAGCTCTGA 1258  
 QY 346 AlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyProSerThrHisIleAsnPheSerIle 365  
 Db 1259 GCAATATGATCTTGCTGCACTCAAGTACTGGGCTCCCTACTACACCAATTTCTCTCG 1318  
 QY 366 GluAsnTyrGlnValGlnLeuGlnLysMetLysAsnMetSerArgGlnGlnTyrValAla 385  
 Db 1319 GAGAAATTATCAGAAAGATTGAGACATGAGAAACATGACATGACAGAAATATTTTCA 1378  
 QY 386 HisLeuArgArgLysSerSerGlyPheSerArgValAlaSerIleTyrArgGlyValThr 405  
 Db 1379 CATTTGAGAGAGAGAGAGAGCTGTTCTTAAGGGGTCTTCATATATGAGAGATCACA 1438  
 QY 406 ArgHisHisGlnHisGlyArgTyrGlnAlaArgGlyLysArgValAlaGlyValAsnLysAsp 425  
 Db 1439 AGACATCACACAGATGAGAGTGGCAAGCAGCATTTGATGAGTGGCTGGAGAAACAAAGAT 1498  
 QY 426 LeuTyrLeuGlyThrPheSerThrGlnGlnGlnValAlaGlyValAlaTyrAspValAla 445  
 Db 1499 CTTTACCTTGGAACCTTTTGAAACCAAGAGAGAGAGCTGCGAAGCTTACATGATGACAGA 1558  
 QY 446 IleLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrAspValGln 465  
 Db 1559 ATTAAGTTCCTGGACAAATGCTGTGACTTAATCTGATATACGAGATGATGATGAT 1618

QY 466 ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGlnLeuValaArgArgLysLysAspAsn 485  
 Db 1619 CGATCATGCTCTAGTAAACACACTCTTGCTGTGAGAGTTAGCGGAAAGAAC----- 1669  
 QY 486 AspProArgAsnLysAspIleAspTyrAsnLysSerValAlaThrSerValAsnGln 505  
 Db 1670 -----AACCAACACCATTTGTC-----GTCCAGAAATACT 1696  
 QY 506 GluThrValGlnValGlnAlaGlyAsnAsnAsnGlnLysAsnAspSerGlnTrpLysMet 525  
 Db 1697 GAA----- 1699  
 QY 526 ValLeuPheAsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLys 545  
 Db 1700 -----GACCAAGACC 1708  
 QY 546 IleMetAsnCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565  
 Db 1709 GCTCTAAATGCT----- 1720  
 QY 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerSerLysIleGly 585  
 Db 1720 ----- 1720  
 QY 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGlnAlaSer 605  
 Db 1721 -----CTTGCGAAGAGCTGCTTCCACAAAGAGATCACT 1753  
 QY 606 ProGlnLysArgGlyProSerLeu-----LeuPheProMetProProMetGluThrLys 623  
 Db 1754 ACTCCGAGAGACTTGTGATTTCCGCGGATTTTGCCTGCTCCATGATATCAAAAG 1813  
 QY 624 IleVal---AsnProIleGlyThrSerValThrSerTrpLeuProSerProThrValGln 642  
 Db 1814 ATGTTGGAATCAAAATATGGCGGAAATATGAGTCTTGACATCAAACTTAATGCTGAG 1873  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTrpThrAsp 662  
 Db 1874 CTTAAG---ACCGTCCTCTTACTTGTGCTCAGATCCGGTTCCTGCTGCGCTGAT 1930  
 QY 663 Thr 663  
 Db 1931 TCT 1933  
 RESULT 7  
 ABX13437  
 ID ABX13437 standard; DNA; 2148 BP.  
 XX  
 AC ABX13437;  
 XX  
 DT 04-JUN-2003 (first entry)  
 XX  
 DE A. thaliana AINTEGUMENTA DNA corresponding to Genbank U40256.  
 XX  
 KW Expression cassette; transgenic; promoter; LOX5; plant; food production;  
 KW animal feed; seed; stress resistance; disease resistance; starch content;  
 KW lipid content; dormancy; fibre content; pharmaceutical production;  
 KW fine chemical production; sterile plant; vitamin; flavouring; perfume;  
 KW dye; cotyledon; embryonic tissues; stress factors; LOX; ds.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN DE10127882-A1.  
 XX  
 PD 12-DEC-2002.  
 XX  
 PF 11-JUN-2001; 2001DE-01027882.  
 XX  
 BR 11-JUN-2001; 2001DE-01027882.  
 XX  
 PA (BADT ) BASF PLANT SCI GMBH.  
 XX  
 PI Bischoff F, Feussner I, Joyall LP;









XX	
PS	Claim 18; Page 508-509; 542pp; English.

The present invention describes an isolated nucleotide fragment (I) comprising a nucleic acid sequence (NS) chosen from a NS encoding a polypeptide (p) having receptor-like protein kinase activity, mitogen-activated protein (MAP)-kinase activity, Ltrf-like transcription factor activity, caldesmon-like activity, ATP citrate lyase activity, SNF1-like activity, and CXC-like transcription factor activity. Also described: (1) complement (II) of (I); (2) a chimeric construct (III) comprising (I) or (II), operably linked to a regulatory sequence; (3) a plant (IV) comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5) oil obtained from (V). (I) or its part can be used in antisense inhibition or co-suppression in a transformed plant. (III) is useful for altering the oil phenotype in a plant such as corn, soybean, wheat, rice, carotol, Brassica, sorghum, sunflower or coconut. (III) is also useful for creating transgenic plants having altered lipid profiles. (I) can also be used as a hybridisation probe. ACC00656 to ACC00868 and ABR40591 to ABR40879 represent sequences used in the exemplification of the present invention.

**Sequence** 1954 BP; 582 A; 379 C; 468 G; 521 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.:	8.53e-83	195
Score:	1104.50	Matches:
Percent Similarity:	52.87%	Conservative:
Best Local Similarity:	40.54	Mismatches:
Query Match:	31.54	Indels:
DB:	7	Gaps:
		27

US-10-024-632-2 (1-663) X ACC00850 (1-1954)

OY		20	GlyPheSerLeuSerProHnISmetyLysMetGlu-----Ala 31
Dd		117	GGGGTCTCTCTCACTCCAGTCCTGAAGATTATGAAAGTTTGAGAAGAAAACCAAGAC 176
OY		32	ThreSerAlaIaThrValProThrThrPheTyMetSerProSerGlnSerIleuSer 51
Dd		177	GTGGAGGAGGGTAGCACCAC----- 194
OY		52	AsnPheGlyMetCysTrpGlyValGlyGluAsnGlyAsnPheHisSerProLeuThrVal 71
Dd		195	--ATCTCTCTCTGT-----CACCTCACTGTCTGT 224
OY		72	MetProLeuLysSerArgGlySerLeuCysIleLeuGluAlaLeuLysArgSer----- 89
Dd		225	ATGCCCTCTGGCTGTATGGCTCTCTCTGTGGTGGACTCTTTAAGCATCGTGGTGA 284
OY		90	----GlnThrGln-ValMetValProThr-----SerSerPr 107
Dd		285	TCTCAAGAAATAAGAAATGTGATATGATGCANATAGGTGGAGGAATTCOAATGAAGAACGTCC 344
OY		100	OlveLeuGluGlyAspPheLeuGlyGluIaThrMetGlyThrHisGluTrpGlySerHisgl 120
Dd		345	AAGGCTTGAAAGACTTTTGGGT----- 366
OY		120	WargGlyLeuSerLeuAspSerIleTrpTyTrpAsnSerGlnAsnIla----- 133
Dd		367	-----TGCCTACTAGACTCCCCGGCTAAAGTCTTTTGGCA 401
OY		136	-GluAlaGlnProAsnArgAspLeuLeuSerGlnProPheArgGlnGlnIyHisMetSe 155
Dd		402	AGACTCTCAAGCTTAACA-----AACCAAAACCAAGAAATGTTCCAAAAATCAA 455
OY		155	IvalGlnThrHisProTyTrpTrpSerGlyLeuAlaCysHisGlyLeuTrpGlnAlaProle 177
Dd		453	TGTTATATGGCACCAAGCTTT-----TGCACCAAC----- 487
OY		175	UcGlnGluGlnThrThrLysGluThrHisValSerAspCysSerSerLeuMetProGlnMe 191
Dd		484	GATTCAAGAGATTGAACCTGGGAGTAAATCTCAAAATCCTCTTCTCTTCTGTCATTTCTT 544

QY	195	-----c-----	ThglGlgIleu-----LysasnTr	202
Db	543	CCATGCAATATATGATTAATCTCAATGCCTTATGCCACCAATGGCATGTAAAGTCTGC	602	
QY	202	pValAlaProThrArgLuphPheSerThrHisGlnGlnValIleuGlnGlnImeLacnScy	222	
Db	603	GTGGGCAACAGACC---CAAGTTTCTTCGTGATGAAAA-----	636	
QY	222	sgIyMetGlyAengIyAaGsnGlyValSerLeuGlySerValGlyCysGlyGluLeuGly	242	
Db	637	----CCTTAATTGAGGCTAATGGGTGCAT-----TTCA	668	
QY	242	nSerIleuSerIleuMetSerProGlySerGlnSerSerCysValThrAlaProSerGly	262	
Db	669	ATCTTGTCCTCCCACTAATGAGCCCAAGTGCATAAATGGG---GTGGGTGCAATTTCTTC	725	
QY	262	yThrAaPseSerValAlaValAaPseAlaLysLysArgGlyHisAlaLysLeuGlyGluIntyGsi	282	
Db	726	TGTTT---CAAGTGAATGAGATAGTAAAGAAACGGGTACATGGCCAAATCTATCTGAGAA	782	
QY	282	nProValHisArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyArgGlyVala	302	
Db	783	ACCAAGTTCCCGGAATCATATGACACTTTTGCGCAAAACAATCTCAATATAGTGTGTGT	842	
QY	302	lThrArgHisArgTrpThrGlyArgTyArgGluAlaHisIleuThrAaPseSerCysLeuGly	322	
Db	843	TACAAAGCAATAGTGTGATCGAAGATAGTAAGAGCCCAATTTGGGATATATGTTCAGAAA	902	
QY	322	sgIyGlyGlnThrArgLysGlyArgGlnValTyLeuGlyGlyTyArgPseMetGluLupy	342	
Db	903	GGAAGGCAAAACAAGAAAGAAAGGCAAGTTTACCTTGCTGTATGATAGAAAGAAAA	962	
QY	342	sAlaAlaAspArgAlaTyArgPseLeuAlaIleuLysTyTrpGlyProSerThrHisIleAs	362	
Db	963	AGAGAGCAAAAGGCTTACGACTTTCAGCTGACCTCAAGATTTGGGGTCCAAACACTCACTAAA	102	
QY	362	nPseSerIleGluAsnTyGlnValGlnLeuGlnGlyImeCysAsnMetSerArgGlnGly	382	
Db	1023	TTTTCCTTTAAGCACTTATGAAAGAAAGACTTGAAGATGAAGCACTGACAGGCAAG	108	
QY	382	cTyTrValAlaHisIleuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyArg	402	
Db	1083	ATTGTGCGCCAAATTGATGAAGAAAGAAAGAGTGAGATTTTCAAGAGGGGCATCTGTGTATG	114	
QY	402	ggIyValAlThrArgHisHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGly	422	
Db	1143	AGGGTGACAAAGCACCATCAACATGAAAGGTGGCAGGCCGCAATAGGAAGGTTGCAGG	120	
QY	422	yAsnLysAspLeuTyLeuGlyThrPheSerThrGlnGlnGluAlaAlaGluAlaTyArg	442	
Db	1203	AAACAAGAACTTGTATCTTCTGTATCACTTTAGCACAAACAAACAACTGCTGAAGCATATGA	126	
QY	442	pValAlaAlaHisIleLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTy	466	
Db	1263	TATAGCGCAATGAAGTTGAGAGAAACAAGTCTGTGCAACTTTGACATTAAGTAAAGTA	132	
QY	462	rAspValGluArgIleMetAlaSerSerAsnLeuLeuAlaGlyLeuAlaArgArgTy	482	
Db	1323	TGATGTGAAGAAGATCTGCTCAAGCTCCACTCTTATTCAGAGAGATCTTGCAAAACGGTC	138	
QY	482	lysAspAsnAspProArgAsnLysAspIle---AspTyArgsn-----LysSerVal	498	
Db	1383	TCTTAAAGAAATCAACAGCTCCGGTTCCTGCTGACTCAATTCGTCGGGTTCATGCC	144	
QY	498	lValThrSerValAsnAsnGlyGluThrValGlnValGlnAlaGlyAsnAsnAsnGly	518	
Db	1443	GATCGCCGTGTGTCTGACGACCTCCTTGGCTATACTGATGGTGAACATTTCTGATGA	150	
QY	518	u---AsnAspSerGlyIuTrpLysMetValLeuPheAsnHisProSerGlnGlnGlnGlnAl	537	
Db	1503	GTTCCTCCAAACATGATGTGG-----	152	
QY	537	aAsnGlyAsnGlySerAspGlnLysIleMetAsnCyGlyAsnTyArgAsnSerAlaPn	557	



QY 220 MetAsnCyseGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGlyCysGly 239  
 DB 752 AACATGAAAGTGGTGCATAGTAGAAGCAGTGTGCAG----- 790  
 QY 240 GluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValThrAla 259  
 DB 791 -----CAGACTCATCACTTCATGATGACTGTTCCACAAACAGC---ACATACATCA 841  
 QY 260 ProSerGlyThrAspSerValAlaValAspAlaLysLysArgGlyLysAlaLysLeuGly 279  
 DB 842 CCCCTTCTCACTGTAGTGTG-----GATATGAGAGAGATTCTTCTGAT 886  
 QY 280 GlnLysGlnPro----- 283  
 DB 887 AACAAAGAACCCATACACGCGCTGCTGATATCAACCCAAACCGAGCCATTGAACCT 946  
 QY 284 ValHisArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValThr 303  
 DB 947 GCACCCAGAAAGTCCATTCATGACACTTTGACAGAGAACTTCATCTACCGTGGTGTAA 1006  
 QY 304 ArgHisArgTyrPheGlyArgTyrGlyAlaHisLeuThrAspAsnSerCysLysLysGln 323  
 DB 1007 AGGATATGCTGGAGCGGGAGGTATGAGGCTCACTGTGGATTAATAGTTGTAGAAAGAG 1066  
 QY 324 GlnGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnGlnLysAla 343  
 DB 1067 GGACAACTCCGCAAGAGAGAGAGAGTTTACTTGGAGAGTTATGACAAAGAGAGAGAGAG 1126  
 QY 344 AlaArgAlaTyrAspLeuAlaLysLysTyrTyrGlyProSerThrHisIleAsnPro 363  
 DB 1127 GCTAGAGCTCATGATTGGCAGCACTTAAATATCGGGAACTGAGCAACAACTTTT 1186  
 QY 364 SerIleGluAsnTyrGlnValGlnLeuGlnGluMetLysAsnMetSerArgGlnGluTyr 383  
 DB 1187 CCAATTAACCACTATGAGAAAGAGTGGAGAAATGAGCACTAGCTAGGCAAGAGTAC 1246  
 QY 384 ValAlaHisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGly 403  
 DB 1247 GTTGGCTCATGAGAGAGAGAGAGTATGCGGTTTCTCGCGGGGATCATTTATCGAGGT 1306  
 QY 404 ValThrArgHisHisGlnHisGlyArgTyrGlnAlaArgTyrGlyArgValAlaGlyAsn 423  
 DB 1307 GTGACGAGACACCATCAATGAGAGAGAGAGAGAGAGATGGAGAGTGTGCGGAC 1366  
 QY 424 LysAspLeuTyrLeuGlyThrPheSerThrGlnGlnGlnAlaGlnAlaTyrAspVal 443  
 DB 1367 AAGGATCTCTACTTGGAACTTTGACACCCAAAGAGAGAGAGAGAGAGATGATGATA 1426  
 QY 444 AlaAlaLysLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrAsp 463  
 DB 1427 GCAGCAATCAATTCAGAGAGAGAGAGTGTCTTCAAACTTTGACATGAGAGATATGAC 1486  
 QY 464 ValGlnAlaGlnIleMetAlaSerSerAsnLeuAlaGlyGlnLeuAlaArgArgLys 483  
 DB 1487 GTGAAAGACATATCTGAGAGACCACTTGTCCAAATGTGTGTCTGCAAAAGCTTTGAG 1546  
 QY 484 Asp-----AsnAspProArgAsnLysAspIleAspTyr 494  
 DB 1547 GATTATGAGAGAGAGAGAGAGAGAGAGATGTTCTATAGAGAGAGATCAAGAGATCAT 1606  
 QY 495 AsnLysSerValValThrSerValAlaAsnAsnGlnGlnThrValGlnValGlnAlaGlyAsn 514  
 DB 1607 ACTAGCATCTGAA--CTCTCACTTAATCAAGAAATCACTTAACAATGAGAGAGAGAG 1665  
 QY 515 AsnAsnAsnGlnLysAspSer--GluTyrLysMetValLeuPheAsnHisProSerGlnGly 534  
 DB 1666 AACACACAGGAGCTCATCATCACTATGAGCAAAAGCTCTGTGATTCACCAACTCAACC 1725  
 QY 534 nglnGlnAlaAsnGlyAsnGlySerAspGlnLysIleLeuLeuGlyAsnTyrArg-- 553  
 DB 1726 TTGCACCACTATCACTACCTTATGACAAAGAAATTAATGTGTGCAAGCAAGACAGA 1785

QY 554 -----AsnSerAlaPheSerMetAlaLeuGlnAspLeuIleGlyIleAspSerValGly 571  
 DB 1786 CAACTCTATGCTCTCTACTTCTTCTTATTCAGATATTCATCACTACAGCTAAGGAA 1845  
 QY 571 YserGlyGlnHisAsnMetLeuAspGluSerSerLysIleGlyThrHisPheSerAsnTh 591  
 DB 1846 CAATGGACACACAACTTCTTACACAAATTCAGAGGTGGACCCCTATGTTAAACATGGA 1905  
 QY 591 rSerSerLeuValThrSerLeuSerSerSerArgGlnAlaSerProGluLysArgGlyPro 611  
 DB 1906 TTCTGCTTCCATGACAACTACTCTTCACTTAACCTCTGTTTATATGATGTATGAGG 1965  
 QY 611 cSerLeuLeuPheProMetProPrometGluThrLysIleVal----- 625  
 DB 1966 TGTGGGGGCTATATGATGATTCCTATGGGAGCTACTACTGTGTGCAATGATGG 2025  
 QY 626 -----AsnPro----- 627  
 DB 2026 TGATCAAAATCCAGAGAGCAATCATGCTTTGGTATATGATTAAGGACATTGGTTA 2085  
 QY 627 ----- 627  
 DB 2086 TGAAGTGTATGTGTCTTCACTGATCTTATCATGACATGCAAGAACTGTATTA 2145  
 QY 628 -----IleGly 629  
 DB 2146 TCTTACTCAACAGCAACATCTTCTGTATGATGACATGAGGCTAGTGCATATGATCAAGG 2205  
 QY 629 YThrSerValThrSerTyrLeuProSer-----ProThrValGlnMetArgProSerPro 647  
 DB 2206 ATCTGATGCAATTAATCTTGGTGTCCACTGCTATTCCAACTCATGACCAAGGTCTAGTAC 2265  
 QY 647 cAlaIleSerLeuSerHis 653  
 DB 2266 TAGTATGGCTCTCTGCCAT 2284  
 DB  
 RESULT 11  
 ACC00801  
 ID ACC00801 standard; cDNA; 2588 BP.  
 AC ACC00801;  
 XX  
 DT 16-MAY-2003 (first entry)  
 XX  
 DE Glycine max oil trait related cDNA sequence SEQ ID NO:351.  
 XX  
 DE Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;  
 KM receptor-like protein kinase; mitogen activated protein kinase; oil;  
 KM lipid-like transcription factor caleosin; ATP citrate lyase; SNF1;  
 KM CKC-like transcription factor; antisense inhibition; co-suppression;  
 KM transgenic plant; gene; ss.  
 XX  
 OS Glycine max.  
 XX  
 PN W02003002751-A2.  
 XX  
 PD 09-JUN-2003.  
 XX  
 PF 27-JUN-2002; 2002WO-US020152.  
 XX  
 PR 29-JUN-2001; 2001US-0301913P.  
 XX  
 PA (DUPO) DU PONT DE NEMOURS & CO E I.  
 XX (PION-) PIONEER HI-BRED INT INC.  
 XX  
 PI Allen SM, Allen WB, Cahoon RE, Epelbaum S, Farnedu OO, Harwell LT,  
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,  
 PI Tarczyński MC;  
 XX  
 DR WPI: 2003-201509/19.  
 DR P-PSDB; ABR40766.  
 XX  
 PT Novel nucleotide fragment encoding polypeptides having receptor-like

protein kinase activity, caleosin-like activity, useful for altering oil phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.

Claim 18, Page 368; 542pp; English.

The present invention describes an isolated nucleotide fragment (I) comprising a nucleic acid sequence (NS) chosen from a NS encoding a polypeptide (P) having receptor-like protein kinase activity, mtogen activated protein (MAP)-kinase activity, LIP15-like transcription factor activity, caleosin-like activity, ATP citrate lyase activity, SNI-like activity and CXC-like transcription factor activity. Also described: (1) complement (II) of (I); (2) a chimeric construct (III) comprising (I) or (II), operably linked to a regulatory sequence; (3) a plant (IV) comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5) oil obtained from (V). (I) or its part can be used in antisense inhibition or co-suppression in a transformed plant. (III) is useful for altering the oil phenotype in a plant such as corn, soybean, wheat, rice, canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for creating transgenic plants having altered lipid profiles. (I) can also be used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to ABR40879 represent sequences used in the exemplification of the present invention.

Sequence 2588 BP; 863 A; 513 C; 554 G; 658 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	Length:	Score:	Matches:
Percent Similarity:	7.09e-78	2588	1048.00
Best Local Similarity:	48.39%	292	Conservative: 84
Query Match:	37.58%	178	Mismatches: 223
	29.93%	223	Indels: 31
DB:	7	Gaps: 31	

US-10-024-632-2 (1-663) x ACC00801 (1-2588)

QY	17	AsntPLeuGlyPhSerLeuSerProHisMetGluAlaThrSer-----Ala 34
DB	42	AAATTGTTAGGTTTCTCTCTCTCTCAGAGAACACCTTTAGTCAAGATCACTCT 101
QY	35	AlaThrValProThrThrPheThrMetSerPro----- 45
DB	102	CAAAAGACACCTTCGTTAGCTTCAACCTGATGATCATCTCAAGACTGATGTA 161
QY	46	-----SerGlnSerHisSerSerPheGlyMet 55
DB	162	GCAGAGAGCTGCTTGAATCTCATCTGACTCAACTCCTCATTTACTTAACTT---CCT 218
QY	56	CysTyrGlyValAlaGlyGluAsnGlyAsnProHis-----SerProLeuThrValMet 72
DB	219	TCTTATGCAATTAAGAA-----GCATTTCACAGAAACATAGATTAACACCACTCAA 272
QY	73	ProLeuLysSerAsp-----GlySerLeuCysIleLeuGlu 84
DB	273	GATTGGAAGAGAGACTTACACACAGCCAAATTTGCTATTGGAAACTTCGTGCAATTAACAA 332
QY	85	AlaLeuLysArgSerGlnThrGlnValMetValProThrSerSerProLysLeuGluAsp 104
DB	333	AAACATGACCAACCAACCAACGCA-----CAGCCAAACCTTGAAAC 374
QY	105	PheLeuGlyGlyAlaThrMetCysThrHisGlyTyrGlySerHisGlyIleuSer 124
DB	375	TTCCTCGGTGGA-----CACTCATTTGGCGAACAATGAG----- 407
QY	125	LeuAspSerIleTyrTyrAsnSerGlnAsnIleGluIleGlnProAsnArgAspLeuLeu 144
DB	407	----- 407
QY	145	SerGlnProPheArgGlnGlnGlyHisMetSerValGlnThrHisProTyrTyrSerGly 164
DB	408	-----CAAAAC-----TACCGTGT 422
QY	165	LeuAlaCysHisGlyLeuTyrGlnAlaProLeuGluGluGluThrThr----- 180

DB	423	AACTCAGCCTTACAGATTACATGTTCTGCTCAGCCAGTATGCGTGTGTGTGT 482
QY	181	-----LysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlu 197
DB	483	AGTGTGTGTGCGAGTACATACATACACACAGTACTCCATAGGATTATCATG----- 536
QY	198	GlyLeuLysAsnTyrPheVal--AlaProThrArgGluPheSerThrHisGlnGlnValLeuG 217
DB	537	-----ATTAAGACATGTTGTGAGAACCAACCCAGAACCTCAGAAACATCA----- 583
QY	217	IuGlnGlnMet--AsnCysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerVal 236
DB	584	-----CAACCAACATGAAAGTGT-----GGCAATTTGAAAGCAGTGTGAG----- 626
QY	237	GlyCysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCys 256
DB	627	-----CAACTCTATCACTTCCATGAGTACTGTTTCAATCAAGC----- 668
QY	257	ValThrAlaProSerGlyThrAspSerValAlaValAlaSerAlaLysArgGlyHisAla 276
DB	659	ACATACCTGCCCCCTTCTCAGCTGTAGTGTG-----GATATGAGAGAGGT 713
QY	277	LysLeuGlyGlnLysGlnPro----- 283
DB	714	TCTTGTGATACAAACCAACCAACCAACCTCGGCTGCACTTATTCACCCAAACCGAGCC 773
QY	284	-----ValHisArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyrArg 300
DB	774	ATTGAACTGCACCCAGAAAGTCCATTGACCTTTGGACAGAAAGACTTCAATCTACCGT 833
QY	301	GlyValThrArgHisArgTyrThrGlyValArgTyrGluAlaHisLeuTyrAspAsnSerCys 320
DB	834	GGTGTAAACAGCCATAGGTGAGCGGAGGCTACAGGCTCACTGTGGATTAATGCTGT 893
QY	321	LysLysGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlu 340
DB	894	AGAGAGAGGAGACACACTCGCAAGAGAAAGCAAGTTTCTGGGTGTTATGATTAAGAA 953
QY	341	GluLysAlaAlaArgAlaTyrAspLeuAlaIleLeuLysTyrTyrGlyProSerThrHis 360
DB	954	GAAGAGGAGCTTAGACCTTACATTTGGCAGCATTAAATCTGGGGAACAACCAACA 1013
QY	361	IleAsnPheSerIleGluAsnTyrGlnValGlnLeuGluGluMetLysMetSerArg 380
DB	1014	ACAAATTTCCATTAGCCATGACATGAGAAAGGTGCAAGAAATGAACACATGACTAGG 1073
QY	381	GlnGluTyrValAlaHisIleLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIle 400
DB	1074	CAAGAGTACGTTGCGTCATGTAGAGAGAGAGTATGAGGTTTCTCGCGGTGATCCATT 1133
QY	401	TyrArgGlyValThrArgHisHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgVal 420
DB	1134	TATCAGAGAGTACAGACAGACACCAACATGAGAAAGTGGCAAGGAGATTGAGAGAGTT 1193
QY	421	AlaGlyAsnLysAspLeuTyrLeuGlyThrPheSerThrGlnGlnGluAlaIleGluAla 440
DB	1194	GCTGGCAACAGAGATCTTTACTTGGAACTTTTACACCCCAAGAGAGGACGAGCAAGGA 1253
QY	441	TyrAspValAlaAlaIleLysPheArgGlyAlaIleAsnAlaValThrAsnPheAspIleSer 460
DB	1254	TATGATGTCAGACATTAATTTCCAGAGACCTAAGGCGCTTACAAACTTTCACATGAGC 1313
QY	461	ArgTyrAspValGluArgIleMetAlaSerSerAsnLeuLeuAlaGlyGluLeuAlaArg 480
DB	1314	AGATATGACGTGAAAGCATACTTGTAGACACACACTTTGCCAATAGGTGTGCTCGAAAG 1373
QY	481	ArgLysLys-----AspAsnAspProArgAsnLysAsp 491
DB	1374	CGTTTGAAGATATGAGACAGGTGTAACGTATGTGATTAATGTCATATGACAGATCA 1433
QY	492	IleAspTyrAsnLysSerValThrSer-----ValSerGlnGlu 505
DB	1434	GTATGATCAT-----AGTATCATCATGAGTTCTTCACTTAATCAAGAAATCAATTAACAC 1487

QY	506	Glut1hrValGlnValGlnAlaGlyAsnAsnSerAsnGluSerAspSerGluTyrIleMet	525
Db	1488	TATGACAGAGGGGGAGAACACGACACATCATATATACGGACATGCTCATGATTCACCA	1547
QY	526	-----ValLeuPheAsnHisPro-----SerGln	533
Db	1548	CCTCAACCTTGACACCCACCATCACTACCTTATGACAAAGATTAATGTCGAAGCAA	16070
QY	534	GlnGlnGlnAlaAsnGlyAsnGly-----SerAspGlnIleLeu	547
Db	1608	GAACACACAGACAACTGTCATGCCCCCTGACTCTTTGTTCTTATGATATTCATCACTT	16670
QY	548	AsnCyseGlyAsnTyrArgAsnSerAlaPhe-----SerMetAlaLeuGlnAspLeu	564
Db	1668	CAGCTAAGGAGCAATGGAACATATACATCTTTTACACAAATGACGGGTTCACCCATAG	17270
QY	565	IleGlyIleAspSerVal-----	570
Db	1728	TTGAGCATGATGATCTGCTTCATTCATGACATATAGCTCTTTCTTAACTCGGTGTTATGAT	17870
QY	571	-----GlySerGlyGlnHisAsnMetLeu-----	578
Db	1788	GGTATATGAGGTGGGGGCTTACATATGTAGTCCCTAGGACATTACTACTGCTGTGTT	18470
QY	579	-----AspGluSerSerIleValIleGlyThrHisPheSerAsnThr-----	591
Db	1848	GCAAGTATGATGTTGATCAAAATTCAGAAAGACATCATGTTGTTGGTGAATATGAGATAAA	19070
QY	592	-----SerSerLeuValThrSerLeuSerSerSerArgGluAlaSerProGlu	607
Db	1908	GCACCTGTTATGAAAGTGTAAGTGTAAGCGCTTCGCACTGATCTTTATCATGACATGCAAG	19670
QY	608	LysArgGlyProSerLeuLeuPheProMetProProMetGluThrIleValAsnPro	627
Db	1968	AAC-----TTGTATATATCTTACTCAACGACATCATCTTCTGTGTGATCA	20120
QY	628	Ile-----GlyThrSerValThrSerTyrLeuProSer-----	638
Db	2013	GTGAAGGCTATGATCATATGATCAAGGCTGTCGATGCAATACTGGGTTCACACTGCTATT	20720
QY	639	ProThrValGlnMetArgProSerProAlaIleSerLeuSerHis	653
Db	2073	CCAACCTCATGCACCGACATCAACTATGATAGCTCTCTGCGCAT	2117
RESULT 12			
ACCC00805			
ID	ACCC00805	standard; cDNA, 1290 BP.	
XX	ACCC00805;		
AC			
XX	16-MAY-2003 (first entry)		
DT			
XX			
DE	Glycine max oil trait related cDNA sequence SEQ ID NO:359.		
XX			
KW	plant; oil trait; oil phenotype; altered lipid profile; MAP kinase; receptor-like protein kinase; mitogen activated protein kinase; oil; LIPID-like transcription factor caleosin; ATP citrate lyase; SNF1; CKC-like transcription factor; antisense inhibition; co-suppression; transgenic plant; gene; ss.		
KW			
XX			
OS	Glycine max.		
XX			
PN	WO2003002751-A2.		
PD	09-JAN-2003.		
XX			
PF	27-JUN-2002; 2002WO-US020152.		
XX			
PR	29-JUN-2001; 2001US-0301913P.		
XX			
FA	(DUPO ) DU PONT DE NEMOURS & CO E I.		
PA	(PION-) PIONEER HI-BRED INT INC.		

PS Claim 18; Page 379-380; 542bp; English.

XX The present invention describes an isolated nucleotide fragment (I)  
CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a  
CC polypeptide (PP) having receptor-like protein kinase activity, mitogen  
CC activated protein (MAP)-kinase activity, lipid-like transcription factor  
CC activity, caldesin-like activity, ATP citrate lyase activity, SMT-like  
CC activity and GGC-like transcription factor activity. Also described: (1)  
CC complement (II) of (1); (2) a chimeric construct (III) comprising (1) or  
CC (11), operably linked to a regulatory sequence; (3) a plant (IV)  
CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)  
CC oil obtained from (V). (II) or its part can be used in antisense  
CC inhibition or co-suppression in a transformed plant. (III) is useful for  
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,  
CC carola, Brassica, sorghum, sunflower or coconut. (III) is also useful for  
CC creating transgenic plants having altered lipid profiles. (1) can also be  
CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to  
CC ABR40879 represent sequences used in the exemplification of the present  
CC invention

XX Sequence 1290 BP; 414 A; 231 C; 321 G; 324 T; 0 U; 0 Other;

SQ

Alignment Scores:

Pred. No.: 4.05e-76 Length: 1290  
Score: 1022.00 Matches: 230  
Percent Similarity: 63.39% Conservative: 47  
Best Local Similarity: 52.63% Mismatches: 90  
Query Match: 29.18% Indels: 71  
DB: 7 Gaps: 12

US-10-024-632-2 (1-663) x ACC00805 (1-1290)

QY 243 SerLeuSerLeuSerMetSerProGlyInserSerCysValThrAlaproSerGly 2626  
Db 6 AGTTGTTCCTCGCAATCTTGAAGCCCCAGTGCGAAATGGG---GGGGTGCATTCTCT 62  
QY 263 ThrAspSerValAlaValAspAlaLysLysArgGlyHisAlaLysLeuGlyGlnVal 2826  
Db 63 GTT---CAAGTAGAATAGAGATAGTAGAAGAACGGGTCATGCCGAATCTCATCTAGAAA 119  
QY 283 ProValHisArgLysSerLLeaspThrPheGlyGlnArgThrSerGlnTYrArgGlyVal 3026  
Db 120 CCAGTTCCTCGCAAAATCTTAACACTTTGGGCAAAACAATCATCATTCGGGGT 179  
QY 303 ThrArgHisArgTrpThrGlyValArgGlyAlaHisLysLeuTPAPAsenSerCysLys 3226  
Db 180 ACAAGGCAATGATGACATCGAAGAATATAGAGGCCCATTTGTGCGATAATAGTTGAGAAAG 229  
QY 323 GluGlyGlnThrArgLysGlyValArgGlnValTyrluGlyGlyTYrAspMetGluGluLys 3426  
Db 240 GAAGGGCAAAACAAGAAAGAAAGGCAAGTTTACCCTGTGTGTATGATAGAGAAAGAAAA 299  
QY 343 AlaAlaArgAlaTyAspSerLeuAlaLysLysTYrTPDjLPFoserThrHisileasn 3626  
Db 300 GCAGCAAAGGCTTACGACTTACCTGCATCANAGTATTGGGATCCAACAACATCACATTAAT 355  
QY 363 PheSerTrleGluAsnTYrGlnValGlnLeuGluGluMetLysAsnMetSerArgGlnGlu 3826  
Db 360 TTCCTTTAGCACTTATAGAGAAGAACTTGAAGAGATGAAGACACATGACAGGCAAGAA 419  
QY 383 TyrValAlaHisLysLysArgLysLysSererGlyPheSerArgGlyAlaIleSerlleTYrArg 4026







Db 2953 GGTGATGAAATGATATGATGTTCTTGATGATCATCTGCTGAATCTTTGGCATGG 3012  
Qy 331 ----- 331  
Db 3013 CTGAGATGATCTTGAGATGACCAAGTACCAAGTACTGGATGATGATGATTTT 3072  
Qy 332 -----ValTyrLeuGlyTyrAspMet 339  
332 -----ValTyrLeuGlyTyrAspMet 339  
3073 GGAACCAAGATTGATTTGATGATTGCAATGTAACCTTTG-CTAGGTGGATGATG 3131  
Qy 340 GluGluValAlaIleArgAlaTyrAspLeuAlaLeuLysTyrTrpGlyProSerThr 359  
3132 GAGAGAAAGCTGCGAGGCGGTATGATCTTGCTGGCTCAAGTACTGGGCGCTTCACG 3191  
Qy 360 HisIleAsnPro----- 363  
3192 CACATCAACTTCCGGTGAATTAATTACCAACTGTGATGATTCATTGCTTTCTT 3251  
Qy 363 ----- 363  
3252 GCAAGCAACAGTACACATTTTGAATTTTGAAGTCTAGCTCAAGGCGACTGTGA 3311  
Qy 364 -----Ser 364  
3312 ATGACTGATGATGATGATTAACGGCTATCATTCGTTGTTTTCACCATTTTCTGGGA 3371  
Qy 365 Ile-GluAsnTyrGlnValGlnLeuGluGlyMetLysAsnMetSerArgGlnIleTyrVal 384  
3372 GTTGAAGACTACCAAGAGAGAGCTGAGAGATGAAGAACATGACGAGAGATGTGT 3431  
Qy 384 LAlaHisLeuArg----- 388  
3432 GGCTCACCCTCGAAGGTACACTGCGTACTGATGAAATCGCAAGTACTGAAAGAA 3491  
Qy 389 -----ArgLysSerSerG 393  
3492 CAAGATTGTGACTGATCATGTGAAACCTCCATGCTGTTGCTGTCGAGAAAGAGCTG 3551  
Qy 393 LysPheSerArgGlyAlaSerIleTyrArgIleValThr----- 405  
3552 GCTTCTCCGCTGGCGCTTCGATCTACGTGAGTCAACAGGTGCTGTATACATGATGT 3611  
Qy 405 ----- 405  
3612 TGACGAAATAATACGCAATGATCTTGAAAGATGTTAACTGATGATTAACCTTGCTG 3671  
Qy 406 -----ArgHisHisGlnHisGlyArgTrpGlnAlaArgIleLeuVal 419  
3672 TGTTTTCTGTCATGTTTGACGATCATTCGACGAGATGACGAGCGGAAATCGGCC 3731  
Qy 419 TgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheSer----- 432  
3732 GGGTCTCGGGCAACAAGACCTTTACTTGGGAGACATTCAGAGTGTCTTCTCCCAA 3791  
Qy 432 ----- 432  
3792 TCTTGCACATGCACTATGATCTACTAGTATGTTCTTGCCGATTCCTTCATCATG 3851  
Qy 432 ----- 432  
3852 TGATCGATCAGCGATGAGCAGCAAGTGTATGAGTTGGCTTTGAGAGGCTGTACG 3911  
Qy 432 ----- 432  
3912 GTTTAAGACGAGGGGTGTCTCTGTCAATGATCATAAAGCCTTACGTAGCTATGCG 3971  
Qy 432 ----- 432  
3972 TTAGATCATGATGCAATTTCTGTGCTCTGACGAGCATGGCCACACATGGCCAG 4031  
Qy 432 ----- 432  
4032 CTCGAAAGAAATGTTTCCCTCCCTGCTTCTTGACATGCTTTTGAATTCGTGACAGCGGTG 4091

Qy 432 ----- 432  
Db 4092 CTTTGGCAGCATGCAATGCGCATGATGCGGGTGTGTTGGTTTGCTTTGTTCG 4151  
Qy 432 ----- 432  
Db 4152 TTCTTTCCAAAGGCTGTGCCCTCGCGCAGTCCGCTGCTTTTTCGCGCGCGCG 4211  
Qy 432 ----- 432  
Db 4212 CGCGCGCCAGCCGGTTGACCTGTGACCTGACCGCGCGCGCGCGCGCGGCGGAAATC 4271  
Qy 432 ----- 432  
4272 GGATTTTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4331  
Qy 433 -----ThrgLingLingLingLingLingLingLingLingLingLingLingLing 449  
4332 TGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 4391  
Qy 449 TGGTAlaAsnAlaValThrAsnProPheAspIleSerArgTyrAspValGluArgIleMet 469  
4392 GGGGCTCAAGCGGCTCAACCACTTCGACATCAGAGGTACGACGTGACCAAGATCTGTG 4451  
Qy 469 LAsrSerAsnLeuLeuAlaGlyLysLeuAlaArgArgLys-----LysAspAsnPro 488  
4452 AGAGCAGACGCTCTCTCCCGGAGAGCTGCGCGCGCGCAAGGTAAAGTCCGACGCG 4511  
Qy 488 TGAAsnLysPheAspIleAspTyrAsnLysSerValThrSerValAsnGlnLysThrVal 508  
4512 GCGGC-----GGGCGCGCGGTGCGCGACCGCGCGCG 4544  
Qy 508 AlGlnValGlnAlaGlyAsnAsnAsnAsnGlnAsnAspSerGluTrpLysValLeu 527  
4545 CTTGTGACGCGCGCG-----AACGTGCGAGAGTGAAGATGCGCACCG 4589  
Qy 528 -----PheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnLysSer 543  
4590 CCGCGCGCTGCGACGCGCGCGCGAGACGACGACGACGACGACGACGACGACGACG 4649  
Qy 543 spGlnLysIleMetAsnCysGlyAsnTyrArgAsnSerValPheSerMetAlaLeuGln 563  
4650 ACCAA-----CACATGACCTCTGCGAGAGCGCTTCTCGGTG-----CTGACAG 4697  
Qy 563 spLeuIleGly-----IleAspSerValGlySerGlyGlnHisAsnMetLeuAspLys 582  
4698 ACATGCTGCGACCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4738  
Qy 582 erLysIleGlyThrHisPheSerAsnThrSerSerLeuValThrSerLysSerSer 602  
4739 -----GCGCGCGACATGTCATGCGCGCG-----ACGAGCTCGGCAACTCC 4781  
Qy 602 TgGluAlaSerProGlnLysArgGlyPro-----SerL 613  
4782 GGGAGCAGAGCCCTGAC-----AGGGGCTGCGCGCGCGCGCGCGCGCGCGCGCTCGCCA 4838  
Qy 613 euleuPheProMetProPheMetGluThrLysIleValAsnProIleGlyThrSerValT 633  
4839 CGCTGTTCGCCAAGCG 4892  
Qy 633 hserTrpLeu-----ProSerProThrValGlnMetArgProSerProAla-----IleSerL 651  
4893 ACACCTGGGCTCGCGCTGCGCGCGCGGTAGCTCGGTCCGCGCGCGCGCGCGCGGTGCA 4952  
Qy 651 euserHisLeuProValPheAlaSerTrpThrAsp 662  
4953 TCGCGACCTCGCAATGTTGCGCGGTGACCGAC 4987  
Db

RESULT 14  
ACCD0799  
ID ACCD0799 standard; cDNA; 1874 BP.  
XX

AC ACC00799;  
 XX 16-MAY-2003 (first entry)  
 XX Glycine max oil trait related cDNA sequence SEQ ID NO:347.  
 DE  
 XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;  
 XX receptor-like protein kinase; mitogen activated protein kinase; oil;  
 XX lipid-like transcription factor caleosin; ATP citrate lyase; SNT1;  
 XX CAC-like transcription factor; antisense inhibition; co-suppression;  
 XX transgenic plant; gene; ss.  
 OS Glycine max.  
 PN WC2003002751-A2.  
 XX  
 XX 09-JAN-2003.  
 XX  
 XX 27-JUN-2002; 2002MO-US020152.  
 XX  
 XX 29-JUN-2001; 2001US-0301913P.  
 XX  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX (PION-) PIONEER HI-BRED INT INC.  
 XX Allen SM, Allen WB, Cahoon RE, Famodu OO, Harvell LT,  
 XX Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,  
 XX Tarczyński MC;  
 XX WPI; 2003-201509/19.  
 XX P-PSDB; ABR40764.  
 XX  
 XX Novel nucleotide fragment encoding polypeptides having receptor-like  
 XX protein kinase activity, caleosin-like activity, useful for altering oil  
 XX phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.  
 XX  
 XX Claim 18; Page 364-365; 542pp; English.  
 XX  
 CC The present invention describes an isolated nucleotide fragment (I)  
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a  
 CC polypeptide (P) having receptor-like protein kinase activity, mitogen  
 CC activated protein kinase (MAP)-kinase activity, lipid-like transcription factor  
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNT1-like  
 CC activity and CAC-like transcription factor activity. Also described: (1)  
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or  
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)  
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)  
 CC oil obtained from (V). (I) or its part can be used in antisense  
 CC inhibition or co-suppression in a transformed plant. (III) is useful for  
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,  
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for  
 CC creating transgenic plants having altered lipid profiles. (I) can also be  
 CC used as a hybridisation probe. ACC00626 to ACC00688 and ABR40591 to  
 CC ABR40819 represent sequences used in the exemplification of the present  
 CC invention  
 CC  
 XX  
 XX Sequence 1874 BP; 542 A; 472 C; 412 G; 448 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 4.75e-70 Length: 1874  
 Score: 952.50 Matches: 248  
 Percent Similarity: 49.64% Conservative: 63  
 Best Local Similarity: 39.74% Mismatches: 153  
 Query Match: 27.20% Indels: 160  
 DB: Gaps: 19  
 US-10-024-632-2 (1-663) x ACC00799 (1-1874)  
 QY 17 AANTPLPLeuGLyPhSeSerLeuSerProHISmetGLyMetGLyAlaThrSerAlaAlaThr 36  
 DB 184 AACTGAGCTTCTCTCTCTCTCTCCCA--ATGGAATGCTCCGACCTCCGAACTCAG 240  
 QY 37 ValProThrThrPheTyrMetSerProSerGlnSerHisLeuSerAsnPhelGlyMetCys 56

DB 241 TTCTTCATATAGACAGCGCGCTTCGGTACTTCTCTCATCTACTACTGCAACTG 300  
 QY 57 TyGlyValAlaGly---GluAsnGlyAsnPro-----HisSerPro 68  
 DB 301 TACACCAACGAGGTGGGCAACGAGCTCAAGTTTGAGCAGATCTCAACCACTGAC 360  
 QY 69 LeuThrValMetProLeuLysSerAspGlySerLeuCysIleLeuGluAlaLeuLysArg 88  
 DB 361 GTGAGTTTGTGA-----TCTTCG 381  
 QY 89 SerGlnThrGlnValMetValProThrSerSerProLysIleLeuGluAsnPhelGly 108  
 DB 382 TCGCAGACGCGCGCCAGACGAGACGACGACCTGCG-----CTGACGACATCTAC 441  
 QY 109 AlathMetGlyThrHisGluTyrGlySerHisGluArgGlyLeuSerLeuAspSerIle 128  
 DB 442 TCTTCGCGC-----GTGATG 456  
 QY 129 TyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsnArgAspLeuLeuSerGlnProPhe 148  
 DB 457 CGTTACTTCGACAGCCAGACGAGACGACGACCTGCG-----CTGACGACATCTAC 510  
 QY 149 ArgGlnGlnGlyHisMetSerValGlnThrHisProTyrTyrSerGlyLeuAlaCysHis 168  
 DB 511 GACCACCAACCAAC-----CACCAACCAACCAACCGCTTACTTCTCTCTCTC 558  
 QY 169 GlyLeuTyrGlnAlaProLeuGluGlu-----GluThrThrLys 181  
 DB 559 GGTGTGACAGCAGAGATCTCAAGCCATTACTGATTTCAAGCTTTTTCGACCACTCC 618  
 QY 182 GluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGluGlyLeuLysAsn 201  
 DB 619 GGTTCGAGGTGATGATTCGATTCATCCATC----- 648  
 QY 202 TrpValAlaProThrArgLupHeserThrHisGlnGlnValLeuGluGlnMetAsn 221  
 DB 649 GGAAGAAGCGAGCGCAGAGATTCGAGCTACTCT-----ATTGAGTCTCTCC----- 696  
 QY 222 CysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGlyCysGlyGluLeu 241  
 DB 697 -----GGCAACGAGATTCGCGCGCTTCGCGGTGACACACCGA----- 735  
 QY 242 GlnSerLeuSerLeuSerMetSerProGlySerGlnSerCysValThrAlaProSer 261  
 DB 736 ---ACCTTGCGCTCGCGCTTGCACTGAGCTCCGAAAGCGCGTTCGCGGAGATCC 792  
 QY 262 GlyThrAspSerValAlaValAspAlaLysValArgGlyHisAlaLysLeuGlnLys 281  
 DB 793 AATAGCTCG----- 801  
 QY 282 GlnProValHisArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyrArgGly 301  
 DB 802 -----AAGAAGTCGTGATCTTCGCGCAGCGACTTCAATTACAGAGGT 849  
 QY 302 ValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTyrAspAsnSerCysLys 321  
 DB 850 GTTACTAGGCAACCGATGACAGAAAGATATGAGGCACTATGAGACATATGTTGCA 909  
 QY 322 LysGluGlyGlnThrArgGlyValArgGlnValTyrLeuGlyGlyTyrAspMetGluGlu 341  
 DB 910 AGGAGAGGTCAAGCCAGAAAGGCGTCAAGTTTATTTGGTGATATGAAGAAGA 969  
 QY 342 LysAlaAlaArgAlaTyrAspLeuAlaLeuLysTyrTyrGlyProSerThrHisIle 361  
 DB 970 AAGCGCGCAGAGCTTATGATTTGGCAGCTTAAGTCTGGGCTCCCACTCTACAC 1029  
 QY 362 AsnPhSerIleGluAsnTyrGlnValGlnLeuGluGluMetLysAsnMetSerArgGln 381  
 DB 1030 AACTTCCTGTTTCATATTCGAAAGAGTGGAGAGATGAACATATCAAGCA 1089  
 QY 382 GluTyrValAlaHisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyr 401

```

Db      1090 GAATTTATTCATCATTCGGGAGAGAAAGTAGTGTTCCTCCAGGGAGCTTCATATAC 1149
Qy      402 ArgGlyValThrArgHisHisGlnHisGlyArgTrpGlnAlaArgLysArgValAla 421
Db      1150 AGAGGTGTTCACAGGCAATCATCAACAGGAGGTGGCAACCAAGTAATGGCCGTGAGCT 1209
Qy      422 GlyAsnLysAspLeuTyrLeuGlyThrPheSerThrGlnGlnGlnAlaGlnAlaTyr 441
Db      1210 GGAACCAAGATTTATCTTGGGAACTTCGCAACCGAGAGAGCAACAGAGGCAATAT 1269
Qy      442 AspValAlaAlaLeuPheAspGlyAlaAsnAlaValThrAsnPheAspLysSerArg 461
Db      1270 GATATTCACACCAATTAAGTTCAAGGTGCAACCGGTAAACCACTTTAGAGTAATAA 1329
Qy      462 TyrAspValGluArgGlyMetAlaSerSerAsnLeuAlaGlyGluLeuAlaArgArg 481
Db      1330 TATGATGTGAAGCTATATATGAAGAGTTCT--CTTCCAGTGGGTGGGAGCAAAACCC 1386
Qy      482 LysLysAspAsnAspProArgAsnLysAspLysPheTyrAsnLysSerValValThrSer 501
Db      1387 TTGAGG----- 1392
Qy      502 ValAsnAsnGlnGluThrValGlnValGlnAlaGlyAsnAsnAsnGlnLysAspSer 521
Db      1393 -----CTTCCCTTGAAATCA 1407
Qy      522 GluTrpLysMetValLeuPheAsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsnGly 541
Db      1408 GAGCAAGAAAGCTCCTCCTGACAGACAGACAGAGTCAAGACAGATCCA----- 1455
Qy      542 SerAspGlnLysIleMetAsnGlyAsnTyrArgAsnSerAla----- 556
Db      1456 -----CAGTGTGGTAACTGAGTGGTGCATCAATTTCTCAGCCATT 1497
Qy      557 -PheSerMetAlaLeuGlnAspLeuIleGlyLysPheSerValGlySerGlyGlnHisAs 576
Db      1498 CATCAAGCCAAATTCCTCAATCCCTTGTG-----GAAATCCG 1533
Qy      576 MetLeuAspGluSerSerLysIleGlyThrHisPheSerAsnThrSerSerLeuValTh 596
Db      1534 TTTCATTCACACACAGCATATTATTCCTCAACCTTTTCAACCTTTTCAACCTTCAAC 1593
Qy      596 rSerLeuSerSerSerArgGlnAlaSerProGlnLysArgGlyProSerLeuLeuPheP 616
Db      1594 GCTGTG-----CAGAGGCTGTGCTTACTTCTGCC 1626
Qy      616 OMePProPro 619
Db      1627 AATGCAACCG 1636

```

```

RESULT 15
ADD30878 standard; cDNA; 1818 BP.
ID ADD30878
XX ADD30878;
XX ADD30878;
XX 15-JAN-2004 (first entry)
XX
XX Plant yield-related polynucleotide clone GI793.
XX
XX de; transcription factor; transgenic plant; growth rate; senescence;
XX seed germination rate; plant vigor; seedling vigor.
XX
XX Arabidopsis thaliana.
XX
XX WO2003013227-A2.
XX
XX 20-FEB-2003.
XX
XX 09-AUG-2002; 2002WO-US025805.
XX
XX 09-AUG-2001; 2001US-0310847P.
XX
XX 19-NOV-2001; 2001US-0336049P.

```

```

PR      11-DEC-2001; 2001US-0338692P.
PR      14-JUN-2002; 2002US-00171468.
XX      (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
XX Ratcliffe O, Riechmann UJ, Adam LJ, Dubell AT, Heard JE;
XX Pilgrim MJ, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
XX Brown PE;
XX
XX MPI; 2003-248221/24.
XX
XX P-PDB; ADD30879.
XX
XX New plant transcription factor polynucleotides and polypeptides, useful
XX in producing transgenic plants with commercially valuable properties,
XX such as an alteration in a plant growth characteristic, e.g. growth rate
XX or apomixis.
XX
XX Disclosure; SEQ ID NO 907; 454bp; English.
XX
XX The invention relates to a number of isolated Arabidopsis thaliana cDNA
XX sequences and their encoded proteins which are especially transcription
XX factor related cDNAs and proteins. The isolated or recombinant plant
XX transcription factor polynucleotides and polypeptides are useful in
XX producing transgenic plants with commercially valuable properties, i.e.
XX modified or altered desirable traits as compared to a reference plant,
XX such as an alteration in a plant growth characteristic, e.g. growth rate,
XX germination rate of seeds, vigor of plants and seedlings, or leaf and
XX flower senescence. Sequence information related to the polynucleotides
XX and polypeptides can also be used in bioinformatic search methods. The
XX transgenic plant is useful for growing a progeny plant from a parent
XX plant. This sequence represents one of the cDNAs of the invention.
XX
XX SQ Sequence 1818 BP; 551 A; 425 C; 414 G; 428 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 5,528-70 Length: 1818
XX Score: 951.50 Matches: 252
XX Percent Similarity: 48.61% Conservative: 80
XX Best Local Similarity: 36.90% Mismatches: 157
XX Query Match: 27.17% Indels: 194
XX DB: 9 Gaps: 23
XX
XX US-10-024-632-2 (1-663) x ADD30878 (1-1818)
XX
XX 14 AsnAsnHisAsnTrpLeuGlyPheSerLeuSerProHisMetLysMetGlnAlaThrSer 33
XX 62 AATCTTAACAACTGGCTTGGCTTCTTCACCGAAC-----AAC 103
XX
XX 34 AlaAlaThrValProThrThrPheTyrMetSerProSerGlnSerHisLeuSerAsn--- 52
XX 104 TCTTCTTGGCTTCCATCAATCAACCTTGGCTTGGTCAGGACCATATGACAAACCT 163
XX
XX 53 -----PheGlyMetCysTyrGlyValGlyGlnAsnGlyAsnPheHisSer 67
XX 164 TTTCACACACAGAGAGGAAATGATCAATCCACAGCGTGAGAGAGAGAT----- 214
XX
XX 68 ProLeuThrValMetProLeuLysSerAspGlySerLeuCysIleLeuGlnAlaLeuLys 87
XX 215 -----GAAGAGGA----- 223
XX
XX 88 ArgSerGlnThrGlnValMetValProThrSerSerProLysLeuGlnAspPheLeuGly 107
XX 224 -----GAGGTTCCAAAGATGGCCGATTTTCTCGGT 253
XX
XX 108 GlyAlaThrMetGlyThrHisGlyLysArgHisGlnArgGlyLysSerLeuAspSer 127
XX 254 -----GTGAGCAAAACCGAGCAAAACCAATCCAACTGAGTACTTCAACACACTTA 307
XX
XX 128 IleTyrTyrAsnSerGlnAsnAlaGlnAlaGlnProAsnArgAspLeuLeuSerGlnPro 147
XX 308 GACTTACTCTTCATCAACCAATAGC---TTGATGCTCT----- 340
XX
XX 148 PheArgGlnGlnGlnLysHisMetSerValGlnThrHisProTyrTyrSerGlyLeuAlaCys 167

```

Db 341 -----ACGCTCCATCAAC----- 355  
 QY 168 HisGlyLeuTyrglnAlaProLeuGluGluThrThrlleGluThrHisValSerAsp 187  
 Db 356 -----GATCGCTTGAGCAGCT 373  
 QY 188 CysSerSerLeuMetProGlnMetThrGluGlyLeuLeuSerThrValAlaProThrArg 207  
 Db 374 TGTGACTCCCAATCTCTTAACAAC----- 397  
 QY 208 GluPheSerThrHisGlnGlnValLeuGluGlnGlnMetAsnCysGlyMetGlyAsnGlu 227  
 Db 398 -----AGTAGCTTCATGAGCTTCAAGAGAGTCTCACAAT----- 433  
 QY 228 ArgAsnGlyValSerLeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeuSer 247  
 Db 434 -----CTACAGTCACTTACTTGTCTC 454  
 QY 248 MetSerProGlySerGlnSerSerCysValThrAlaProSerGlyThrAspSerValAla 267  
 Db 455 ATGGGACCAACCGCTGTAAATATGTTGTAGACAAAGCTTCAACATCCGAGACCCGG 514  
 QY 268 ValAspAlaValSerGlyAsnGlyHisAlaValSerLeuGlyGlnLeuProValHisArgLys 287  
 Db 515 GATAACCTAGCGGTGAGAGCACTAGCCGTGTGAGACGGCCACGCCA-----AGACGT 568  
 QY 288 SerLeuAspThrPheGlyGlnArgThrSerGlnThrArgGlyValThrArgHisArgLys 307  
 Db 569 GCATTGACACTTTCGACAAACGAACTCATCTGCTGTGTCGACAAAGACATCAATGG 628  
 QY 308 ThrGlyArgTyrglnAlaHisLeuThrAspAsnSerCysLeuGlyGlnGlnThrArg 327  
 Db 629 ACTGCTGATGTAGGCTCATCTATGGGATATAGTTGTAGAGGGAAGCCAGCTCAG 688  
 QY 328 LysGlyArgGlnValTyrgLeuGlyGlyTyraAspMetGluGluValAlaAlaGlyAlaTy 347  
 Db 689 AAAGAGAAACAAGTTTACTTGGGTGATGACAAAGAAAGATTAAGACGAAATCATATA 748  
 QY 348 AspLeuAlaAlaLeuTyrgTyrgGlyProSerThrHisLeuAsnPheSerLeuGluAsn 367  
 Db 749 GATCTAGCTGACCTTAAGTACTGGGTCTTCAACTCTACTAATTTCCCATTTACAAC 808  
 QY 368 TyrglnValGlnLeuGluMetLysAsnMetSerArgGlnGluTyrgValAlaHisLeu 387  
 Db 809 TACGAGAAAGAGTAGAGAAATGAGACATGACGAGACAGAGTTCGTGGCTGCCATT 868  
 QY 388 ArgArgLysSerSerGlyPheSerArgGlyValAspSerLeuTyrgArgGlyValThrArgHis 407  
 Db 869 AGAAGAGAAAGTAGTGGATTTCAGAGAGCCCTTCAGATGATCGAGAGATTACAAGCAT 928  
 QY 408 HisGlnHisGlyArgTyrgGlnAlaArgLysGlyValAlaGlyAsnLysAspLeuTyrg 427  
 Db 929 CACCAACATGGAAGTGGCAAGCAAGATCGCCGAGTCCGGGAAACAAAGACCTCTAC 988  
 QY 428 LeuGlyThrPheSerThrGlnGluGlnAlaGlnAlaGlnAlaTyrgAspValAlaAlaLeuLys 447  
 Db 989 TTGGGAACCTTTAGACTGAGAGAGAGACAGACAGCTTACGATATAGCTGCATTAAG 1048  
 QY 448 PheArgGlyAlaAsnAlaValThrAsnPheAspLysSerArgTyrgAspValGluArgGly 467  
 Db 1049 TTTAGAGAGCTTAATGCAAGTACCACTTCAGATCAACCGGTACGAGCGTGAAGGCATT 1108  
 QY 468 MetAlaSerSerAsnLeu-----LeuAlaGlyLeuLeuAlaArgArgLysAsp----- 484  
 Db 1109 CTAGAGAGTAGACTTTCCTCATCGAGAGCGCGACCTTAAACGGCTCAAGAGCTCAA 1168  
 QY 485 -----AsnAspProArgAsnLysAspLysAsp----- 493  
 Db 1169 GCTCTTAGAGTCTTCAAGAGAAACGAGAGCGAGATGATAGCCCTTGTTCAAGTTTCCAG 1228  
 QY 493 ----- 493

Db 1229 TACGCTGTGCTGAGACAGAGCTCTGCTCCACCTCATCAAGACTTCAAGCTTCAACCT 1288  
 QY 494 TyrAsnLysSerVal-----ValThrSerValAsnAsnGluThr 507  
 Db 1289 TACCTCTTAAGCATTTCAACCAACATTAAGAGCTTTTCTATCTCTTCAAGAACATGAC-- 1345  
 QY 508 ValGlnValGlnAlaGlyAsnAsnAsnAsnGluLeuAsnAspSerGlyThrPlyMetValLeu 527  
 Db 1346 -----ATCTCTCATTAACAACAACAATCTCAACATCTCTCC-----TCT 1387  
 QY 528 PheAsnHisProSerGln----- 533  
 Db 1388 TTTAATCAACCAATCAATATATCAACAACACTTCACTTCCACCAACAGACCAATTAAC 1447  
 QY 534 -----GlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLysIleMetAsnCysGlyAsnTy 552  
 Db 1448 TTGCGACAAAGTGTAGCCAGAAC-----TCTACAGCTCTACATGCTATCTCAT 1501  
 QY 553 ArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIleGlyIleAspSerValGlySer 572  
 Db 1502 AGCAATCCGCTCTG-----CTTCATGACTGTCTCTTACCTCATGCTGACACAC 1532  
 QY 573 GlyGlnHisAsnMetLeuAspGluSerSerLysIleGlyThrHisPheSerAsnThrSe 592  
 Db 1553 AATAATTAACAATGAGAGCTTAGTGGAGCTTACAAACACTGACAGCATTTCTTGGAAACAC 1612  
 QY 592 rSerLeuValThrSer-----LeuSerSerSerArgGluLysAspProGluTy 608  
 Db 1613 GGTATTGTATTTGGTCCAGCTCGACTGTGTGATCGACCGAGGAGTTCCAAACGTTAAA 1672  
 QY 608 SerGlyProSerLeuLeuPheProMetProPheMetGluThrLys-----IleValAs 626  
 Db 1673 ACGG-----ATTACGATATAGCTTCCAGTATGGAAACGAGGGATATATGTTGG 1723  
 QY 626 nProfile 628  
 Db 1724 ACCAATG 1730

Search completed: March 13, 2004, 05:51:51  
 Job time : 486 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 13, 2004, 00:44:08 / Search time 4685 Seconds  
(without alignments)  
6133.710 Million cell updates/sec

Title: US-10-024-632-2

Sequence: 1 MKXINESNTDGNHNLG.....RPSAISLHLPVFSWTD 663

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=x1h  
-Q=/cgm2\_1/USPTO.spool/US10024632/rnuc.09032004.101207.10765/app.query.fasta.1.839  
-DB=Genbank -QFMT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blcsm62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pct -NORM=exc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10024632.ccm 1.1.2860 @runat.09032004.101207.10765 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESUBERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WASN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

GenBank: 1: gb\_ha: 2: gb\_hgt: 3: gb\_hn: 4: gb\_om: 5: gb\_ov: 6: gb\_pac: 7: gb\_ph: 8: gb\_pl: 9: gb\_pt: 10: gb\_ro: 11: gb\_scs: 12: gb\_sy: 13: gb\_un: 14: gb\_vl: 15: em\_da: 16: em\_fun: 17: em\_hum: 18: em\_in: 19: em\_mu: 20: em\_om: 21: em\_or: 22: em\_ov: 23: em\_pac: 24: em\_ph: 25: em\_pl: 26: em\_ro: 27: em\_scs: 28: em\_un:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3502	100.0	2344	6	AX555216 Sequence
2	1906	54.4	2546	8	AY461432 Nicotiana
3	1715	49.0	2323	6	AX555218 Sequence
4	1408.5	40.2	1926	6	AX555220 Sequence
5	1408.5	40.2	2510	8	AK106306 Oryza sat
6	1349.5	38.5	2010	6	AX555223 Sequence
7	1328	37.9	1699	8	AY117207 Arabidops
8	1328	37.9	1905	8	ATU41339 Arabidops
9	1328	37.9	2056	8	AY080706 Arabidops
10	1328	37.9	2148	6	BD274516 Methods f
11	1328	37.9	2148	6	AR316367 Sequence
12	1328	37.9	2148	6	AR427901 Sequence
13	1328	37.9	2148	6	ATU40256 Arabidops
14	1328	37.9	1738	6	BD274518 Methods f
15	1307	37.3	1738	6	AR316369 Sequence
16	1307	37.3	1738	6	AR427903 Sequence
17	1307	37.3	1738	6	ATT28119 Arabidops
18	1095.5	31.3	110766	8	ATTC8198 Arabidops
19	1095.5	31.3	198493	8	ATTC8198 Arabidops
20	977	27.9	152684	2	AX555222 Sequence
21	977	27.9	152684	2	AX555222 Sequence
22	949.5	27.1	1440	8	AC133450 Oryza sat
23	947	27.0	1992	8	AF317905 Brascica
24	947	27.0	2011	6	AX058689 Sequence
25	944	27.0	2014	6	AX058687 Sequence
26	944	27.0	2014	6	AR317904 Brascica
27	933	26.6	123820	8	AC103891 Oryza sat
28	933	26.6	166072	8	AC135157 Oryza sat
29	914	26.1	1745	8	AK101959 Oryza sat
30	890	25.4	1773	8	ZMWHCPI 247554 Z.mays mRNA
31	876	25.0	2153	8	AK111891 Oryza sat
32	850.5	24.3	1650	8	AY062179 Oryza sat
33	762	22.8	1560	8	AK109839 Oryza sat
34	713	20.4	133739	8	AP005309 Oryza sat
35	695	19.8	121675	2	AC144645 Medicago
36	672	19.2	1069	8	AY133880 Arabidops
37	672	19.2	1237	8	AY045915 Arabidops
38	668	19.1	1572	8	AK111617 Oryza sat
39	664.5	19.0	1866	8	AK106769 Oryza sat
40	664.5	19.0	1910	8	AK112088 Oryza sat
41	653	18.6	5151	6	AX058692 Sequence
42	653	18.6	5151	6	AR317907 Arabidops
43	647.5	18.5	33563	8	AT11086 Arabidops
44	647	18.5	81677	8	AB025629 Arabidops
45	630	18.0	105807	8	AC006085 Arabidops

RESULT 1

## ALIGNMENTS

AX555216  
 LOCUS AX555216 2344 bp DNA linear PAT 27-NOV-2002  
 DEFINITION Sequence 1 from Patent WO02059332.  
 ACCESSION AX555216  
 VERSION AX555216.1 GI:25898744  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Glycine max (soybean)  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 REFERENCE  
 1 He, S.S. and Dotson, S.B.  
 Nucleic acid molecules associated with plant cell proliferation and  
 growth and uses thereof  
 Patent: WO 02059332-A 1 01-AUG-2002;  
 Monsanto Technology LLC (US)  
 FEATURES  
 source  
 1..2344  
 /organism="Glycine max"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3847"  
 242..2233  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD58225.1"  
 /db\_xref="GI:25898745"  
 /translation="MKRIINSEINNTDGDGNHMLGFSLSPHMKKEATSAATVPTFPYKS  
 PSOSLSNFGMCGYGENGNHSPLETPVMDGLSLILELKRSGOTQVWVPSPL  
 EDPLGGATMGTHEYSGHERGLSLDSIYVNSONARAONRDLQSPFGCGHMEVQHP  
 YVSGIACHGILYQAPLEBETTKETRYSDCSSLMPQMTGLKMWAPTEESTHQVLEQ  
 QNCGMGNKRGVSLGVSYGCELSLSLSPGSCVTPASGTSVAADAKRQVLA  
 KLQKQPVHRSIDTFQGRISQYRGVTRHWGTRYEAHLMDNSCKEGQKRRQVLA  
 GGDMEEKARAADLAALKVGPSTHINFSLENYQVLEEMKMSRQEVYAHIRKSS  
 GFSRGASLYRGVTRHHQGRWQARIGRVAQKDLVLTGTFSTQEEAAYVAALKRPS  
 ANATNFDISRYDVERIVASNLGELARLKKNDNRNDIDYKNSVIVSNEETV  
 QVQAGNNNNDEEMKMYLFNHPGQQQANNGNSDOKIMNGVNRNSAFPMALQDLIG  
 IDVSGGQHNMLDESSKICHTFNSMTSLYLSRSSRSLSPKRPGLLFPMPMETKI  
 VNPIGTSTSWLPSPFTVQMRPSPAISLSHLPFASWTDY"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,02e-228 Length: 2344  
 Score: 3502.00 Matches: 663  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0  
 US-10-024-632-2 (1-663) x AX555216 (1-2344)

QY 1 Metlysaagllleasgluserasasanthraspaeqlyaaenaahiasettpleugly 20  
 DB 242 ATGAAGCCCATTAATGAAGATPAACAACCATGATGAAACAATCAATCGTTGGG 301  
 QY 21 PheSerLeuSerPhePheIleMetLeuMetGluAlaThrSerAlaAlaThrValProThrThr 40  
 DB 302 TTCTCTCTCAACCCCATGAAATGAGGCTACTTACAGCAGCAGCTGTTCCGACAAC 361  
 QY 41 PheTyrMetSerProSerGlnSerHisLeuSerAsnPheGlyMetCysTyrGlyValGly 60  
 DB 362 TTCAACAGTCCCTCTTCAATCTCACTGTCACAACTTCGGAAGTGTCTACGGGTGGA 421  
 QY 61 GluAsnGlyAsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeu 80  
 DB 422 GAAATGGAATCTTCAATCTTCCACTTACGGTTATAGCCTCTCAAGTCCGATGGTCTCT 481  
 QY 81 CysIleLeuGluAlaLeuLysArgSerGlnThrGlnValMetValProThrSerSerPro 100  
 DB 482 TGTATCTTGGAAAGCTCTCAAAAAGATCAACAACGCAAGATGATGTCACAACTTCCTCCG 541

QY 101 LysLeuGluAspPheLeuGlyGlyAlaThrMetGlyThrHisGlyTyrGlySerHisGly 120  
 DB 542 AATATGAGAGCATTTCTTAGTGTGTGTCACATAGGGAATCTACAGATATAGAACCAACGAG 601  
 QY 121 ArgGlyLeuSerLeuAspSerIleTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsn 140  
 DB 602 AGAGGTTTACGCTTACAGACATCTATATATATATATATATATATATATATATATATATAT 661  
 QY 141 ArgAspLeuLeuSerGlnProPheArgGlnGlnGlyHisMetSerValGlnThrHisPro 160  
 DB 662 AGAGACCTTCTTCAACACCTTCAAGCACAAGGTCATATAGTGTCCAAACACACCTT 721  
 QY 161 TyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGluGluGluThr 180  
 DB 722 TATTACTCAGGCTTGTGCTTGCATGTTATATATATATATATATATATATATATATATATATAT 781  
 QY 181 LysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGluGlyLeuLys 200  
 DB 782 AAGAAACGACAGCTGTGCGATTCAGCTTCTTAATGCTCAATATACAGAGGCTTGAA 841  
 QY 201 AsnTPValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGluGlnMet 220  
 DB 842 AACTGGGTGCTCCAAACAAGGAGTTTCAACTCACAGCAGGTTTGGAGCAGCAATG 901  
 QY 221 AsnCysGlyMetGlyAsnGluLysArgGlnValSerLeuGlySerValGlyCysGlyGlu 240  
 DB 902 AATTGGCGATGGGAGATGAGAAATGATGTTGTTTATGAGATCTGTGGGTGTGGAAG 961  
 QY 241 LeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValThrAlaPro 260  
 DB 962 TTACAGTCTTAAGCTTATCTATGATCTGTGTTTCTAGTCTAGTGTGTCTCTCTCT 1021  
 QY 261 SerGlyThrAspSerValAlaValAlaSerAlaLysLysArgGlyHisAlaLysLeuGlyGln 280  
 DB 1022 TCTGGAAACGATTCCTGTTGCTGTGATGATCAAGAAAGAAAGGCAATGATGATGATG 1081  
 QY 281 LysGlnProValHisArgGlySerSerIleAspThrPheGlyGlnArgThrSerGlnTyrArg 300  
 DB 1082 AAGCACTCTGTCTATGAAATCTTATCAACATTTGGGCAAGAACTTCGACATATG 1141  
 QY 301 GlyValThrArgHisArgGlyTrpThrGlyArgGlyAlaHisLeuTyrAspAsnSerCys 320  
 DB 1142 GGTGTCAACAAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1201  
 QY 321 LysLysGluGlyGlnThrArgLysGlyValArgGlnValTyrLeuGlyGlyTyrAspMetGlu 340  
 DB 1202 AAGAAAGAGGCAAACTAGAAAGAGCAAGAGCAAGAGGATTTGGGGGATTAATGATGAG 1261  
 QY 341 GluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyProSerThrHis 360  
 DB 1262 GAGAAAGCTGCAAGAGCTATGATCTCGGGCCCTTAAGTACTGGGAGCACTTCAACGAT 1321  
 QY 361 IleAsnPheSerIleGluAsnTyrGlnValGlnLeuGluGluMetLysAsnMetSerArg 380  
 DB 1322 ATAACTTTTCATAGAGATTAACAAGTCAACTTGAAGAAATGAACAACCTGGACGA 1381  
 QY 381 GlnGlyTyrValAlaHisLeuArgArgLysSerSerGlyPheSerArgGlyValAsnIle 400  
 DB 1382 CAGGATAGCTTGACACCTTGAGAAAGAAAGCAAGCGGGTTTTCAGAGTCTCAATA 1441  
 QY 401 TyrArgGlyValThrArgHisHisGlnHisGlyArgTyrGlnAlaArgIleGlyValArgVal 420  
 DB 1442 TACAGAGGGGTACAAAGCATCAACAATGAAATGCAAGCAAGATGAGGCAAGT 1501  
 QY 421 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheSerThrGlnGluAlaAlaGluAla 440  
 DB 1502 GCTGGAAACAAGACCTTACCTTGGAAGCTTGACACCAAGAGAGAGAGCAGAGACCA 1561  
 QY 441 TyrAspValAlaAlaIleLysPheArgGlyValAsnAlaValThrAsnPheAspIleSer 460  
 DB 1562 TACGATGTAGGGCGCATCAAAATTTGCGCGCCCAATGACATCAAACTTTGACATTTCA 1621  
 QY 461 ArgTyrAspValGluArgIleMetAlaSerSerAsnLeuAlaGlyGluLeuAlaArg 480



```

Db      1622 AGATCGATGTGAGAGATCATCGCCAGTACCAATCTCTCCCTGGGAGCTTGACAGG 1661
Qy      481 AAglySLYASpASpASpProArghAsnLYSAspIleAspTYRAsnLYSAspValValThr 500
Db      1682 CGTAAGAAAGATACGATCTCTAGAAACAAGACATAGACTACCAACAGAGTGTAGTACA 1741
Qy      501 SerValAsnAsnGluThrValGlnValGlnValAsnAsnAsnGluAsnAsp 520
Db      1742 AGTGGACAAATGAGGAAAGGCTCAAGTCAAGCAAGAAACAACAAATTAAGAAACAC 1801
Qy      521 SerGluTPLYMeValLeuPheAsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsn 540
Db      1802 TCAGAGTGGAGATGTTTATTTATTAACACCCCTCAGACGACGACAAAGCAAAATGGCAAT 1861
Qy      541 GYSerAspGlnLYLeuMetAsnGlyAsnTYRAsnSerValAsnSerValAspSerMetAla 560
Db      1862 GGACATGACCAAAATATATGAACTGTGAAATTACAAACAGTGCATTTTCTATGACC 1921
Qy      561 LeuGlnAspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlu 580
Db      1922 CTACAGAGATCTTATTGGGATTGATTGGTGGGTTCTGGGACGCAATATCTGTGACGAG 1981
Qy      581 SerSerLYLeuGlyThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSer 600
Db      1982 TCTACCAAAATGGGACTCATTTTCAAAACAGTCATGCTGCGTGCACAAAGTTTAAGCAGC 2041
Qy      601 SerATGGLuAlaSerProGluLYSArgLYProSerLeuLeuPheProMetProProMet 620
Db      2042 TCAAGAGAGGCTTAGCTGCTGAGAAAGGGTCCCTGCTCTGTTCCCAATCTCCCAAGT 2101
Qy      621 GlnThrLYSLeuValAsnProIleGlyThrSerValThrSerTPLYProSerProThr 640
Db      2102 GAAACAAAGATTGTCAACCCATGTGTACAGTGTACTTGTGGTCACTCAACCAAG 2161
Qy      641 ValGlnMetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTyr 660
Db      2162 GTTCAAAAGAGGCTTCTCTGCTATCTCTTGTCTCACTTGCACAGTTTGTCTTGG 2221
Qy      661 ThrAspThr 663
Db      2222 ACTGATCT 2230

RESULT 2
LOCUS      AY461432 2546 bp mRNA linear PLN 29-NOV-2003
DEFINITION Nicotiana tabacum ANT-like protein mRNA, complete cds.
ACCESSION  AY461432
VERSION     AY461432.1 GI:38492171
KEYWORDS
SOURCE
ORGANISM   Nicotiana tabacum (common tobacco)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
REFERENCE  1 (bases 1 to 2546)
AUTHORS   Rieu, I. and Mariani, T.
TITLE      A tobacco ANT homolog is expressed in the whole ovary
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 2546)
AUTHORS   Rieu, I. and Mariani, T.
TITLE      Direct Submission
JOURNAL    Submitted (07-NOV-2003) Plant Cell Biology, University of Nijmegen,
Toernooiveld 1, Nijmegen 6525 ED, The Netherlands
FEATURES
Source
1..2546
/organism="Nicotiana tabacum"
/mol_type="mRNA"
/cultivar="petit havana SR1"
/db_xref="taxon:4097"
343..2274
/note="similar to Arabidopsis thaliana ANT protein"
/codon_start=1

```

```

/product="ANT-like protein"
/protein_id="AA02338.1"
/db_xref="GI:38492172"
/translation="MKMSKMDNNKNSHRSSNNNNNNNSHNDASNWLQFSASP
HKMEVTSSEPOHQQFNNOSSALAAQSYLSSPNVATSNLSALCYGVGNPPGHS
SLSVPLKSDGSLCTIMEALSRSAAAMQSSSPKLEPFGASQVGNERRAMALSL
DELVYHONEGYYSPMCHGMTQESLIDETKRPIDISCDQMTAVSNEKSKGHVATD
OHNETCSSWVGAGGGGTSVGGGVGNDLQSLSNPDSQSCVPRQISFVEL
ECVALETKRBSAGKVVQKQPVARKSIDTFGQRTQYRGVTRHRTGYSRAHLMDSCK
KEQVYALRRKSSGFSRGSAMRYGTVRQGRMORLIGRVANQOLYGTGSTOEA
AEAYDVAAIKRPVNAVTVTFDLSRVDEVTIMSNLIPAGELARRNERRERPDYD
KNEAEACVQNDNNNNNNNTVDMKMSVSSSSLSIGSRVRRPTSMALQDIDYDMS
NRIIDDDQNKICGNFNSASLVTSIGSSREASPKSAASLVFAFPTVNACTPSAQL
RIPVYSISLHPVPAALNDA"

ALIGNMENT Scores:
Pred. No.: 5.5e-120 Length: 2546
Score: 1906.00 Matches: 405
Percent Similarity: 68.88 Conservative: 73
Best Local Similarity: 58.368 Mismatches: 114
Query Match: 54.438 Indels: 102
DB: 8 Gaps: 21

US-10-024-632-2 (1-663) x AY461432 (1-2546)

Qy      3 ArgIleAsnGluSerAsnSerHisPheAspGlyAsnAsnHis-----Asn 17
Db      391 AGAAGTACAAATATATATATATATATATATATATATAGTCAATAGTACTAGTAACTAC 450
Qy      18 TPLYGlyPheSerLeuSerProHisMetLYSMeGluAlaThrSer----- 33
Db      451 TGGTTAGATTTCCTCTCCACCCACATGAAAGAGAGTTACTTCTCTGAACCTCA 510
Qy      34 -----AlaAlaThrValProThrThrPheTyrMetSer 44
Db      511 CACCAACATCATGAGTCAATATATACAGACTGCTGCTGCTCAAAAGTTCTATCTTCA 570
Qy      45 ProSerGlnSerHisLeuSerAsn-----PheGlyMetCysTyrGlyValGluSer 62
Db      571 TCTTCTCCATGATAGTGAATCTCAATACCTGCTGCTGTATAGGTTGAGAGAAAT 630
Qy      63 GlyAsnPhe--HisSerProLeuThrValMetProLeuLYSAspGlySerLeuCys 81
Db      631 AATCCCTTGGCCATCTCTTATCTGTATAGCTCTCAATCTGATGGCTCTTGGC 690
Qy      82 IleLeuGluAlaLeuLYSArgSerGlnThrGlnValMetVal---ProThrSerSerPro 100
Db      691 ATTATGGAAGCTCTCTTCANAGTCTCATGACAGATGCTATGGGCAAAAGTCTCTTCACT 750
Qy      101 LysLeuGluAspPheLeuGlyValAlaThrMetGlyThrHisGluTyrGlySerHisGlu 120
Db      751 AAGCTTGAAGACTTTTGTAGTGTGCAAGT-----CAATATGGAATCATGAA 798
Qy      121 Arg-----GlyLeuSerLeuAspSerIleTyrTyrAsnSerGlnAsnAlaGluAla 137
Db      799 AAGGAAGCTATAGCTCTTAAGCTTACAGCTTATATATATAC----- 840
Qy      138 GlnProAsnArgAspLeuLeuSerGlnProPheArgGlnGlnGlyHisMetSerValGln 157
Db      841 -----CAA 843
Qy      158 ThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGluGlu 177
Db      844 AATGAAGCTATATATCTCCCATGATGATGATGATGATGATGATGATGATGATGATGAT 903
Qy      178 GlnThrThrLYSLeuThrHisValSerAspCysSerSerLeuMetProGlnMetThr--- 196
Db      904 GAA---ACAAACCAACATCAATTTCAATGATGATGATGATGATGATGATGATGATGATGAT 960
Qy      197 GlnGlyLeuLYSAsnTyrValAlaProThrArgGluPheSerThrHisGlnGluValLeu 216

```

Db 961 AATGAAGTAAAGTTGG-----GGGCTTATGCTATGACCAACTATATAAT 1008  
 Qy 217 Glu-----GlnGlnMetAcnGlyMetGlyAsnGluAsnGlyValSerLeu 233  
 Db 1009 GAAATCTTAGATAGATAGTGTGTGTCAGAGTGTGTGGAGTACTCTGCTGTGTGTGT 1068  
 Qy 224 GlySerValGlyGlyGlyGlnLeuGlnSerLeuSerLeuSerMetSerProGlySerGln 253  
 Db 1069 GGTACAGTTGGTGTATGATTACAGTTGAGCTTGTGATGAAATCCGTGTTCTCAG 1128  
 Qy 254 SerSerCysValThr-----AlaProSerGlyThrAspSerValAlaValAsp 269  
 Db 1129 TCTACTGTGTATTACCAAGCAAAATTTCTCTATAGCTGAGATGTTGCTCATTTGA 1188  
 Qy 270 AlaIleValArgGlyHisAlaLeuGlnGlnProValHisArgGlySerIle 289  
 Db 1189 ACAAGAAAGAGGCTTCTGAAAGAGTTGTTCAAAACCACTGTTCTATAGGAATCTATT 1248  
 Qy 290 AspThrPheGlyGlnArgThrSerGlnThrArgGlyValThrArgHisArgThrPheGly 309  
 Db 1249 CATACATTTGGTCAAGAACTTCTCAGATAGAGTGTCACTAGGCAATAGTGAACAGT 1308  
 Qy 310 ArgTyrGlnAlaHisLeuTyrAspAsnSerCysIleGlyGlnThrArgGlyGly 329  
 Db 1309 AGGTATGAGGCGCATTTGTGGATATAGTTGCAAGAGAGGCGAGACTAGAAAGGA 1368  
 Qy 330 ArgGlnValTyrLeuGlyGlyTyrAspMetGlnGluValAlaIleArgValTyrAspLeu 349  
 Db 1369 AGCGAAGTTTATCTTGGGGGCTATGATATGAAAGAGAAAGCTGCAAGACATATGATTTA 1428  
 Qy 350 AlaAlaLeuIleTyrTyrGlyProSerThrHisIleAsnPheSerIleGluAsnTyrGln 369  
 Db 1429 GCTGCACTAAAGTATTTGGGAGCTTCACTCACTTAATCTTCCTGATAGAACTTACAG 1488  
 Qy 370 ValGlnLeuGlnGlnMetLeuAsnMetSerArgGlnGlnTyrValAlaHisLeuArgArg 389  
 Db 1489 AAAGAGCTTGAAGATATGAAATATGACCCGGCAAGATATGTTGCACTTATAGAGG 1548  
 Qy 390 LysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisIleGln 409  
 Db 1549 AAAATAGTCAATCTTCAAGAGAGTCTTCAATGATATGAGAGAGCAAGCACTATCAG 1608  
 Qy 410 HisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGly 429  
 Db 1609 CATGGAAGATGGCAGCAAGAAATGGAGAGTGTGAGGAAACAAGATCTTCACTTGGG 1668  
 Qy 430 ThrPheSerThrGlnGlnGlnAlaGlnAlaTyrAspValAlaIleIleLysPheArg 449  
 Db 1669 ACTTTCAGCACACAAAGAAAGCTGCAAGACATATGATGTTGCTGCATTAATTCAGA 1728  
 Qy 450 GlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrAspValGlnValGlnIleMetAla 469  
 Db 1729 GGTGTAATGTCTGACAAACTTGTGACATATCGCATCGACGTGAGAAATCATGCT 1788  
 Qy 470 SerSerLeuLeuAlaGlyGlnLeuAlaArgArgLysLysAspAsnProArgAsn 489  
 Db 1789 ACTATATACACTTCCGTCTGAGAAATTAAGTACGACAAACAAGAAAGAGACCA----- 1842  
 Qy 490 LysAspIleAspTyrAsnLysSerValThrSerValAsnAsnGlnGlnThrValGln 509  
 Db 1843 -----ATTGATATACATCTCAGT-----GATCAACAAGATGAAAGAGCACTGTCTCAA 1890  
 Qy 510 ValGlnAlaGlyAsnAsnAsnAsnGlnAsnAspSerGlnTyrLysMetValLeuPheAsn 529  
 Db 1891 AATGACACACACACACACATGGCAATATGTCACAGATTTGAAAAATGCTCATAT--- 1947  
 Qy 530 HisProSerGlnGlnGlnAlaAsnGlnLysAsnGlySerAspGlnLysIleMetAsnCys 549  
 Db 1948 -----CAATCCCAAT-----CCTTCTCTA 1968  
 Qy 550 GlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIleGlyIleAspSer 569  
 Db 1969 GGAATATATAGAAACCAACTTTTCCATGGCATTACAAGATCTATATGCTATTTGATTG 2028

Qy 570 ValGlySerGlyGlnHisAsnMetLeu---AspGlnSerSerValIleGlyThrHisPhe 588  
 Db 2029 ATGAATTCACAGCAACCAATTCATCTTATGATGATCAACAAACAGATGATGTAACATTTT 2088  
 Qy 589 SerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGlnAlaSerProGlnLys 608  
 Db 2089 TCAATGCTTCATCTTGTGTTGCTTCCAGCTTGGCACTTCAAGAGAACCAAGTCTTATATA 2148  
 Qy 609 ArgGlyProSerLeuLeuPheProMetProPoweMetGlnThrLysIleValAsnProIle 628  
 Db 2149 AGTCAGCCTCTTGTGCTTTCGCAAGCCA----- 2178  
 Qy 629 GlyThrSerValThrSerTyrLeuProSerProThrValGlnMetArgProSerProAla 648  
 Db 2179 ---ACAAATGTCAAATGATATTCCTCT-----GCTCAACTAGGCAATTCACA--- 2226  
 Qy 649 IleSerLeuSerHisLeuProValPheAlaSerTyrThrAsp 662  
 Db 2227 GTTTCATTTCTCACCTACCACTTCTTGTCTGCTTGAATGAT 2268  
 RESULT 3  
 AX555218  
 LOCUS AX555218 2323 bp DNA linear PAT 27-NOV-2002  
 DEFINITION Sequence 3 from Patent WO02059332.  
 ACCESSION AX555218  
 VERSION AX555218.1 GI:25898746  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Glycine max (soybean)  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 REFERENCE  
 1 He, S. S. and Detson, S. B.  
 Nucleic acid molecules associated with plant cell proliferation and  
 growth and uses thereof  
 Patent: WO 02059332-A 3 01-AUG-2002;  
 Technology LLC (US)  
 Monsanto  
 FEATURES  
 source  
 1..2323  
 /organism="Glycine max"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3847"  
 25..2022  
 /note="unassigned protein product"  
 /codon\_start=1  
 /protein\_id="CAD58226.1"  
 /db\_xref="GI:25898747"  
 /db\_xref="RMTREMBL:CAD58226"  
 /translation="MKSMDNDNADLNQNNWLGFSLSPOKHNGICVSSHQSSAAEV  
 VPTSEYVHTAPLSSYGYGLAEENVGLYSALPIPLKSDSLVLETLRSQAQAMA  
 TSTPELNFLEGEAMGTPHNYECATMPLESLDSVFIQPSRDPNNNOYQNHVQ  
 HISTNQOQOQQLQAYSTLRNHWMLTEGSKOSQSDNNNLHVQMGDDAPVYGLK  
 SVEVNFQSHHSEKMLVPHYEENAGSGSIGVAYGLDGLSLSMSPSQSSSVTS  
 SHRAAYDAVDAMTKRGPEKDKQKQIVHRKSTDTGCGSTQYRGYTRRWGTGRYE  
 ALHMDNSCKEGRQSKRGQVYIGGIDMEKARAPADIALAKTWGSTHTINPFLYQYQ  
 ELBEMKNTRQSYVAHLRKSGSFRGASMRGVTNRHOGHWQNRKRVAGNKKVLYL  
 GFSTQSEAAEAYDIALIKPRGANAVTFDITRYDVEKIMASNNLSSELARNRKED  
 NTKQCIDQHNKPSAYEDTOEALIMHQKSCSENDQMVLYYOSQOLEQNPFTESD  
 RNQCFEVALDMMFHQVEESKATHTVSNPSLSLSSSRGSPDRTSLPMLSGMP  
 STASKLTLNPNVNVMSWDPSPHLRPAITLIPQVPFAAWTDA"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,55e-107 Length: 2323  
 Score: 1715.00 Matches: 378  
 Percent Similarity: 65.86% Conservative: 83  
 Best Local Similarity: 54.00% Mismatches: 165  
 Query Match: 48.97% Indels: 74  
 DB: 6 Gaps: 20

US-10-024-632-2 (1-663) X AX555218 (1-2323)

Oy	MetValSerAlaLeuSerAsnAsnThrAspArgLysAsnAsnHisAsnTrpLeuGly	20
Db	ATGAAGAGATATGAGAAATGATGACATCTGACTTATTAATCAAAACAATGGGTGGT	84
Oy	21 PheSerLeuSerProHisMet-----LysMetGluAlaThr	32
Db	TTCTACACTCTCTCCCAATGATATATAGAGTTTCTTCACTCAACACTCTCCCT	144
Oy	33 SerAlaAlaIleValProThrThrPheTyrosMetSerProSerGlnSerHisIleuSerAsn	52
Db	GCTGCTGAAGTGCTTCTTCAACAGCTTTTATCCACCACACTGCTCCA-----CTTAAGTAGC	198
Oy	53 PheGlyMetCysTrpGlyVal---GlyGluAsnGlyAsnPheHisSerProLeuThrVal	71
Db	TATGGTTTCTACTATGAGACTTGAACTGAAATGTGGATGTATCTTACGCTTGGCAATC	258
Oy	72 MetProLeuLysSerAspGlySerLeuGlySerIleLeuGluAlaLeuLysArgSerGlnThr	91
Db	ATGCCCTCCAAATCTGATGGCTCTCTCTATGATGATTGGAACTTTAAGAGGTCAACAACA	318
Oy	92 GlnValMetValProThrSerSerProIlyseuGluAspPheLeuGlyGlyAlaThrMet	111
Db	CAAGCAATGGCTACTACTTCAACACCAAACTGGGAGAACTTTGTGGTGGGAGGCCCATG	378
Oy	112 GlyThr---HisGluTyr-----GlySerHisGluArgGlyLysSerLeuAspSer	127
Db	GGGACCCCTCATCATTAAGATATGATGTCACACAGAAACAAGTCCCTGATGACTTAGACGT	438
Oy	128 IleTyTrpAsnSerGlnAsnAlaGluAlaGlnProAsnArgAspLeuLeuSerGlnPro	147
Db	GTTTATTAC---ATCCACCCCTCAGCCGTGACCCCAATATATACCAAACTTACCAAAAC	495
Oy	148 PheArgGln-----GlnGlyHisMetSerValGlnThrHisIleProTyTrp	162
Db	CATGTTCAACACATTATAGCACCAACAACAACAACAAGACACTTCAAGCATATATAC	555
Oy	163 SerGlyLeuAlaCysHisGlyLeuTyrglnAlaProLeuGluGluGlnThrThryGlu	182
Db	TTTACCTTGGAGAAACCATGATATGATATATTA-----GAAGGGCTCAAAAGCAA	600
Oy	183 ThrHisValSerAspCysSerSerIleuMetProGln-----	194
Db	AGCCAAACTTCTGACACACACATCTTCATGTTCAAAACATGGGTGATGATGCTGCTT	660
Oy	195 MetThrGluGlyLeuLysAsnTrpValAlaProThrArgGluPheSer-----Thr	211
Db	CCGTGTTCCGGCTCAGAGGTGG-----GAAGTGAAGAACTTCAAGCTAGCCATCA	714
Oy	212 HisGlnGln-----ValLeuGluGlnGlnMetAsnCysGlyMetGlyAsnGluArg	228
Db	CATGAGCTCAAAAGATGATGTTCTTCATGTCAGAGAAATGCTGCGTGAATACAGG-----	768
Oy	229 AsnGlyValSerLeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeuSerMet	248
Db	-----TTCATATGATCAATGCTTATGTGTGACTTGCATTCGTATGAGCTTGCCATG	819
Oy	249 SerProGlySerGlnSerSerCysValThr-----AlaProSerGlyThr	263
Db	AGTCTTAGCTCTCACTATGACAGTGTACCAAGTTCACCGGCTTCAACTGCTGCTGCT	879
Oy	264 AspSerValAlaValAspAlaLysIlySerArgGlyHisAlaLysLeuGlyGlnLysGlnPro	283
Db	GATTCGTGTCATGATGATCTAAGAAAGGGGGCTGAAGAGTTGACCAAGAAAGAAATT	939
Oy	284 ValHisArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyTrpArgGlyValThr	303
Db	GTTCAATGAGAGTCCATTCATGACACTTTGGACAAAGAACTCCCAAGTATAGAGATTAACA	999
Oy	304 ArgHisArgTrpThrGlyArgTyrglyAlaHisIleuMetTrpAspAsnSerCysIlySerGlu	323
Db	AGGCATATGGTGAATGGAGATATGAAGCTCATCTTTGGGACAAACAGCTGCAAGAAAG	1055

ACCESSION AX555220  
 VERSION AX555220.1 GI:25898748  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
 AUTHORS He, S.S. and Dobson, S.B.  
 TITLE Nucleic acid molecules associated with plant cell proliferation and  
 growth and uses thereof  
 JOURNAL Patent: WO 0205932-A 5 01-AUG-2002;  
 Monsanto Technology LLC (US)  
 FEATURES  
 source  
 1..1926  
 /organism="Oryza sativa"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4530"  
 1..1926  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD58227.1"  
 /db\_xref="GI:25898748"  
 /translation="VASSGGSSNNMIGFSLSPHMPAMEVSPSEPTAAHHHHHPRA  
 AAAAGASSPPDASTTCNPI.PSPRAAMVA.PSPGYVVGAYGDSGTAGVYSHLP  
 VMPISDGLCLMEGMSPSSPKLEDFICGNGSDPAITYSGGCEADASSAYOH  
 HQLVYNTQPLTEAMLEDAAPMEDMAAKNPLTYSACTGEMPOPLSLS  
 PGOSSSCVSAAPQHQAVVAALAAAGDQSNNDGSEQRVGRGTVGSGQKQ  
 VRRKSIDTFGRTSQYRGVTRHRTGRYEALMDNSCKDQTKRQVLYGVDTE  
 KARAYDLAALKYMLSTHINPLENYDEIEEMRMRQRYVAHLRRSSGFRGAS  
 IYRGVTRHQRHQRWQARIGRVANKDILGTSTOEAAEYDIALIKFGILAVTAR  
 DITRDVDKIMSSSLPGEARVKYALEAPDPHPIGRLEGATREASADIVPGKISG  
 NVLHSGQQQALACTEATRADLOKMGMRHSHALNGTGYVESAAADIVPGKISG  
 IYFNSNSLIVTSLSNRSRSGPERGLAHLAKHHPITAVSLAAMPKMPAPAAAHVM  
 RPPSAIAHL.PVPAAWTDA"

ORIGIN

Alignment Scores:  
 pred. No.: 2 31e-86 Length: 1926  
 Score: 1408.50 Matches: 340  
 Percent Similarity: 55.058 Conservative: 63  
 Best Local Similarity: 46.458 Mismatches: 152  
 Query Match: 40.224 Indels: 177  
 DB: 6 Gaps: 26

US-10-024-632-2 (1-663) x AX555220 (1-1926)

QY 13 GYAASHNHIASNTPLLEUlyPheSerLeuSerProHisMetLeuMetGluAlaThr 32  
 Db 13 GGGGAGGAGGACAACTGTTAGCTTCTGCTCCCGCACATGCGGCATGAGAGGTG 72  
 QY 33 SerAlaAlaThrValProThrThr----- 40  
 Db 73 CCGTCTCTCTGAGGACATGCTGCTCATCATCATCATCATCATCATCATCTGCT 132  
 QY 41 -----Phe 41  
 Db 133 GCTGCTGCTGCTGCGGAGCATATGCTCTCTCCGACAGGCGACGACCTGCAACTTC 192  
 QY 42 TYMetSerProSerGlnSerHisLeu-----SerAspHeGly 54  
 Db 193 CTTTCTCTCTCTCTGACACAGATGCTGCTCTTCACTGCTACTACTACTGCTGGC 252  
 QY 55 MetCysTyrGlyValGlyLysn-----GlyAspPheHisSerProLeuThrValMet 72  
 Db 253 GGGGCTTACGAGACCGGACACGACCGCGGCTTACTACTCTGCACTCCCTGTCATG 312  
 QY 73 ProLeuLysSerAspGlySerLeuCysIleLeuGluAlaLeuLysArgSerGlnThrGln 92  
 Db 313 CCAATCAAGTCCATGCTCTCTCTCATCATGAAAGC----- 351

QY 93 ValMetValProThrSerSerProLysLeuGluLysPheLeuGlyAlaThrMetGly 112  
 Db 352 ---ATGATGCGCTGCTCATCGCCAAAGCTTCATGGAGTGGGCAATGGCAGT 408  
 QY 113 ThrHisGluTyrGlySerHisGluLysArgLysSerLeuAspSerIleTyrAsnSer 132  
 Db 409 GGCAT-----GACCCGCGCACTACTATAGC 435  
 QY 133 GlnAsnAlaGluAlaGlnProAsnArgAspLeuLeuSerGlnProPheArgGlnGlnLys 152  
 Db 436 CAGGCGCAAGAGCAGAGAT-----GCAAGCAGGCGGCTTACCAAGAC 480  
 QY 153 HisMetSerValGlnThrHisIleProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGln 172  
 Db 481 CACCACTAGTCT-----CCCTAACAC-----TACCAAG 507  
 QY 173 AlaProLeuGluGluGluThrThrLysGluThrHisValSerAspCysSerSerLeuMet 192  
 Db 508 CCATTGACCGAAGCAGAGATGCTGCAAGCGCGGCGGCAATGAGAGCAATG 567  
 QY 193 ProGlnMetThrGluGlyLeuLysAsnTyrValAlaProThrArgGluPheSerThrHis 212  
 Db 568 -----CGCGCGGCGCAAGACTTC----- 585  
 QY 213 GlnGlnValLeuGluGlnGlnMetAsnCysGlyMetGlyAsnGluLysArgGlyValSer 232  
 Db 586 -----CTGCTCACACGACTACGCGGCTCTACGCAACACAGAGATGCCG----- 630  
 QY 233 LeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySer 252  
 Db 631 -----CAGCCCTCAGCCTTCCATGACCCAGGCTCC 663  
 QY 253 Gln---SerSerCysValThr---AlaProSerGlyThrAspSerValAlaValAspAla 270  
 Db 664 CAGTCAGACAGCTGCTCATGTCAGCTCCCGACAGCATCAGCAGATGCGGCTGCTGCT 723  
 QY 270 ----- 270  
 Db 724 GCAGCTGCTGCTGCTGATGACGAGGAGCAACAGTATGACGCTGCGGACGCGT 783  
 QY 271 -----LysLeuArgGlyHisIleAlaLysLeuGlyGlnLysGlnProValHisArgLysSer 288  
 Db 784 GTCGGAGAGAGAGGAGGACCGGAGAAAGGAGGCAAAAGCAGCTGTTACCGGAGATTC 843  
 QY 289 IleAspThrPheGlyGlnLysArgHisSerGlnTyrArgGlyValThrArgHisArgTyrThr 308  
 Db 844 ATTGACACGTTTGGGCAAGAGACATCGCAGTATAGGGGCTGACCGACAGGAGTGGACT 903  
 QY 309 GlyArgTyrGluAlaHisIleuTyrAspAsnSerCysLysLeuGluGlyGlnThrArgLys 328  
 Db 904 GGAAGATATGAAACCCACTCTGGGATACAGTTGCAAAAGATGGAACAGCAAGAGAG 963  
 QY 329 GlyArgGlnValTyrLeuGlyLysTyrAspMetGluLysGluAlaAlaArgAlaTyrAsp 348  
 Db 964 GGAAGCAAGTATATAGTATAGTATGACATGAGAAAGACCTCGAGGCTTATGAT 1023  
 QY 349 LeuAlaAlaLeuLysTyrTyrGlyProSerThrHisIleAsnPheSerIleGluLysTyr 368  
 Db 1024 CTGGCTGCTGTAATACTGGGGGCTATCTTACCAATTAATTTCCCGTTAGAAATCTAC 1083  
 QY 369 GlnValGlnLeuGluGluMetLysAsnMetSerArgGlnGluTyrValAlaHisIleuArg 388  
 Db 1084 CCAAGATGAGATCGAAGAGATGGAAGATGACAGCAAGAAATATGTCGCACTTGAGA 1143  
 QY 389 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHis 408  
 Db 1144 AGAGAGAGCAGCGGATCTCTCGCGTCTTCCATTTACCGGAGATACAGGCTATC 1203  
 QY 409 GlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeu 428  
 Db 1204 CACATGAGAGATGAGCAGCTGCGATGCGAGGCTGCGCAACAGCAGCATTTGATCTC 1263  
 QY 429 GlyThrPheSerThrGlnGluGluAlaGluAlaTyrAspValAlaIleLysPhe 448

[illegible]

TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL	Science 301 (5631), 376-379 (2003)
MEDLINE	22752273
PUBMED	12869764
REFERENCE	2 (bases 1 to 2510)
AUTHORS	Adachi, J., Aizawa, S., Akimura, T., Arikawa, T., Carninci, P., Doi, K., Fujiwara, T., Fukuda, S., Hanaoka, T., Hara, A., Hashizume, N.,

TITLE  
JOURNAL  
Direct Submission  
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of  
Agrobiological Sciences, Department of Molecular Genetics, Head of  
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki  
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,  
Tel:81-29-838-7007, Fax:81-29-838-7007)  
This clone is one of the 28K full-length cDNA clones from japonica  
rice.

FEATURES  
 source  
 /organization="Oryza sativa [japonica cultivar-group]"  
 1. 2510  
 Location/Qualifiers  
 NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,  
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J.,  
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,  
 Onmeda, E., Yanagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and  
 Yamamoto, M.  
 PAIS Genome Sequencing & Analysis Group: Otsuno, Y., Iida, Y.,  
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
 Mizuno, K., Naitikawa, R., Nitkawa, J., Oka, M., Ryu, R., Sugano, S.,  
 Sugiyama, A., Suzuki, Y., Tsunoda, J., Ueda, M., Xie, Q., Yokomizo, S.,  
 Yoshimura, A., Matsubara, K. and Murakami, K.  
 Genome Exploration Research Group in Riken: Adachi, J., Aizawa, K.,  
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hangaki, T.,  
 Hara, A., Hasehime, M., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
 Hirooka, T., Hori, F., Iida, J., Imamura, K., Imocani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,  
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Komoh, H., Kouda, M.,  
 Koyas, S., Kunihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Nunabe, K., Ohno, M., Oosato, N.,  
 Oota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
 Takaki-Akibira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K.,  
 Yasunishi, A. and Hayashizaki, Y.

```

/mol_type="mrna"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="002-101-Cl0"

```

```
/ad_xref="Faxon:3994/"
```

```
/clone="002-101-C
```

### Alignment Scores:

Pred. No.:	3.3e-86	Length:	2510
Score:	1408.50	Matches:	340
Percent Similarity:	55.05%	Conservative:	63
Best Local Similarity:	46.45%	Mismatches:	152
Query Match:	40.22%	Indels:	177

Score:	1408.50	Matches:	340
--------	---------	----------	-----

Percent Similarity:	55.05%	Conservative:	63
---------------------	--------	---------------	----

Best local similarity:	46.45%	Mismatch:	152
------------------------	--------	-----------	-----

DEBC LOCAL SIMILARLY:	70:128
Query Match:	40:228
INDEXES:	177

Query Marcin:	40.228	Index:	171
DB:	0	Gars:	26

US-10-024-632-2 (1-663) X AK106306 (1-2510)

QY	13	G1YAsnSnnIAsnTPRLysGIuPhleSerLeuSerProH1MeLysMetGluAlaThr	32
Db	194	GGGGGAGAGAGAACAGTGGTATGGCTTCTGGCTCTCCCGGACATCGCGGCAATGGAGGTG	253
QY	33	SerialAlaThrValProThrThr	40
Db	254	CCGGTCCTCCCTGAGCCATGAGCTGGCTGCTCATCATCATCATCATCATCATCT	312
QY	41	-----	41
Db	314	GGCTGCTCTGCTCCCGGAGCCATGCTGCTCTCCCGGACAGGCGGACAGCTGAACTTC	372
QY	42	TyrMetSerProSerGlnSerHisLeu	54
Db	374	CTCTCTCCCTCCCTCGAGACAGAGATGGTCTCTTCACTGGCTACTACTAGCTGGC	433
QY	55	MetCysArgIglValGlyIleuAsn-----GlyAsnPheHisSerProLeuThrValMet	72
Db	434	GGGGCTACGAGAGCGGGAGACAGACACCGCGGGGTACTACTACCATCTCCGGTATG	492
QY	73	ProLeuLysSerArgIglLysLeuLysCysG1IleLeuGluAlaLeuLysArgSerGlnThrGln	92
Db	494	CCATACAGTCCGACAGGCTCTCTCCATCATATGGAGGC-----	533
QY	93	ValMetValProThrSerSerProLysLeuGluAspPheLeuGlyIglAlaThrMetGly	112
Db	533	---ATGATGGCGTGCATGCTGCGCAAGCTCGAGAACTTCTGGGGTGTCGGAATGGCAGT	582
QY	113	ThrHisGluLysGlySerHisGluAspGlyLeuSerLeuAspSerIleTyrTyrAsnSer	133
Db	590	GGGCAT-----GACCGGCGCACTTACTATGCG	612
QY	133	GlnAsnAlaGluAlaGlnProAsnArgPheLeuSerGlnProPheArgGlnGlnGly	152
Db	617	CAGGGCCAAAGACAGAGAT-----GCCAGAGAGGCGGCTTACACAGAC	665
QY	153	HisMetSerValGlnThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGln	172
Db	662	CACCAAGTAACTC-----CCCTCAAC-----TACGAG	682
QY	173	AlaProLeuGlnGluGlnThrThrIlysgIuThrHisValSerAspCysSerSerLeuMet	192
Db	689	CCATTGACGGAACAGAGATGCTGCAGAGGCGCGGCGGCGGCAATGAGAGCGCAATG	742
QY	193	ProGlnMetThrGluGlyLeuLysAsnTPRLValAlaProThrArgGluPheSerThrHis	212
Db	749	-----GGGGCGGCGCAAGAACTTC-----	762
QY	213	GlnGlnValLeuGlnGlnGlnMetAsnCysGlyMetGlyAsnGluAlaArgAsnGlyValSer	232
Db	767	-----CTGCTTACACAGCTACGCGGCTGCTGACGGCAACAGAAATGCG-----	811
QY	233	LeuGlySerValGlyCysGlyGluLeuGlnInsertLeuSerLeuMetSerProGlySer	253
Db	812	-----CAGCGCTACGCTTTCATATAGGCCAGGGCTC	844
QY	253	Gln---SerSerCysValThr---AlaProSerGlyThrAspSerValAlaValAspAla	272

Db	845	CAGTCAGACAGCTGGTCAAGTCAGACGCTCCACAGAGATCAGACAGATGACGAGTGGCGGTGGCT	270
Qy	270	-----	270
Db	905	GCAGCTGCTGCTGCTGATGCGCACGAGAACAGATATGACGCTGGCAGACGCT	964
Qy	271	-----LYSLYSARGLYVHIALALYLLEUGLYGLILYSGINPRGYALHISARGLYSER	288
Db	965	GTCGGGAAAGAGAGGGGCACCGGAAAGGGGGCCAAAAGCAGCTGTTCACCGGAAATCC	1022
Qy	289	ILASPETHRPHGLYGLIARGTHSRSEGLTYTRARGLYVALTHRYGHSARGTPPTTR	308
Db	1025	ATTGACACGTTGGGCACAGGACATCCGAGTATGAGGGCGCTCACCGAGCACAGTGGACT	1082
Qy	309	GLYARGLYTGLUALAHILEUTIPASPSANSECYSLYSYLSGLIUGLYGLINTHRYGLYS	328
Db	1085	GGAAAGATAGAGCCCACTCTGGGATATACGTTGCAAAAAGATGACAGACAAAGAG	1144
Qy	329	GLYARGGLNVALTYRLEUGLYGLITYRASPMEGLIUGLYLVALALARGALATYRSP	348
Db	1145	GGAGGCAAGTATATCTGATGSGTTATGACCTGAGATATAACTCCGAGGCTTATGAT	1202
Qy	349	LEUALAHILEULYSYTYTRIGLYPROSEETHRIEILEANPSESETRIEGLIUSNTYR	368
Db	1205	CTGGCTGGCTAAATATCTGGGGGCTATCTACGATATTAATTTCCCGTTAGAAAATCC	1262
Qy	369	GLNVALGLNLEUGLIUGLUMETLYASPMSESETRIGLIUGLYVALAHISLEULYR	388
Db	1265	CGAGATGATGAGAGAGATGGAAGATGACAAAGGAGAAATATTTGCGACTTGAGA	1322
Qy	389	ARGLYSESESEGLYPHESETRARGLYVALSERILETYRZRGLYVALTHRYGHSIS	408
Db	1325	AGGAGAGACAGCGGGTCTCTCGCGGCTTCATCTACCGGGAGATTACAGAGCATAC	1382
Qy	409	GLNHISGLYARGTTPGINALARGILGLYARGVALALGLYASNLYSAPLEUTYLEU	428
Db	1385	CAGCATGAGAGATGGCAAGCTGGATTTGGCGGGTTGCTGGCAACAAGACTTGATTC	1442
Qy	429	GLYTRPHESETRHGLIUGLIUALAAGLIUALATYRSPVALAALATLEYSPE	448
Db	1445	GGCACTTCAGACACTCAAGAGAAGACGACAGAGCATACACACTTCTCCATCAAGATTC	1504
Qy	449	ARGGLYALASNLIALVALTHRASPAPASPLESETRGYRSPYALIGLINTGLILEUC	468
Db	1505	CGTGGCTGGAACCGGTGACGAACTTTGACATCCACAGGTGACGACTGGACAGATCATG	1562
Qy	469	ALASERSEASNLLEULAGLYGLIUEUALATYRGLYSLEYS-----	483
Db	1565	GAGAGACAGCTCCTGCTGCTGATGAGGACGCGGTAGGTGAAGCGATCGACGAGCGG	1622
Qy	484	-----ASPASPAPRPGARGASNLYSAPRIILASPTYRASNLYSESERVALALHRSERVAL	502
Db	1625	CCGACACATGTCCTCAATAGCCCGCGAGCTC-----	1662
Qy	503	ASPAENGLIUGLTHRVALIGLIVALIGLINALGLYASNAASNAENGLIUSNAPSERGILU	522
Db	1661	ACCGAGGAAGCAGCGGCTGCTACTGATCACGGGC-----ACCGAC	1692
Qy	523	TRPLYSMETVALILEUPHAEENHISPROSEEGINLGLIUALA-----ASN	538
Db	1700	TGGAGATGCTGCTC-----CARGATCAACACGACAGCAAGCTGCAAGCTGCACACGA	1752
Qy	539	GLYASNGLYSERAPRGLIYSLIEMETASNCYGLYASNTYRARGASERVALAPHER	558
Db	1754	GCAAGCGCAGATCTTCAGAG-----GGCTTCATGGGTGACCGCGCACTCG	1792
Qy	559	METALALEUGINAPLEULIEGLYILEASPSEVALIGLYSERGLYLNHISMETLEU	578
Db	1799	-----GGCTTCAACGGCATTTGGGTTGACGTGAGTGGCGGCGACGTCACAGATCGAT	1852
Qy	579	APRGLUSERSERLYILE-----GLYTHHISPHESERHANTHSERSEULEVALHISER	597







```

QY 345 rgaiaayrappieualalaletuetyrtiprglyproserthrhislleampheseri 365
Db 1055 gggggaatgattcttgcctccctcaagracgagcccttccagacacattcccg 1114
QY 365 legiuanyrgrlvalgluclueclueclueclueclueclueclueclueclue 385
Db 1115 tgcagagactacacagagagagctgagagagatgagagagacagagagatgag 1174
QY 385 lahlaleuargrlysersegrlypheserarglyalaserlletyraglyval 405
Db 1175 ctacactcagaaagaaagacagcttccctccctccctccctccctccctcc 1234
QY 405 hrarghahleaglhiaglyrgrlvalaarglleglyagvalalaglyasulysa 425
Db 1235 ccagagatcatcagacagagagatgagagagacagagatgagagagacagag 1294
QY 425 spleutyrtleuglythpthe----- 431
Db 1295 acccttacttggagacattcattccgcttgccttgccttgccttgccttgcct 1354
QY 432 --serthrglglucluaalaglualaayrappvalalalalelysephargly 451
Db 1355 ccggacacagagagagagagagagagagagagagagagagagagagagag 1414
QY 451 laernaalavalthrasnpheserlyserarglyrappvalgluargllewerlaser 471
Db 1415 tcacagccgctacacacacttgcacatcagagagatgagagagatgagagag 1474
QY 471 efasnleuclualaglylueclualagarglyls--lyaspsasnaspproargan 490
Db 1475 gacacctctctccggggagagctggcgccgagagagagagagagagagagag 1533
QY 490 yasaplleasprlyasnylservalaltnrservalasnasnclugluthrvalglu 510
Db 1534 -----gcggcgccgctgctggcgccgagcgccgagcgccgcttgg 1567
QY 510 aiglialaglyasnasnleuclualasnaspsarglytrlysemetvalleu----- 527
Db 1568 tgcagagccggg-----acgtgcggagatgagagatgagagagagagag 1612
QY 528 -----pheasnhleproserglnglgluclualasnglyasnglysearpo 545
Db 1613 ccctgcagagcgccgagagagagagagagagagagagagagagagagagag 1671
QY 545 yalllewerasncyseglasnyrargasnaseralphasermetalaleuclue 565
Db 1672 -----cacatgaccttccctgcagagagagagagagagagagagag 1720
QY 565 leglly--lleaspservalgllysearglylghisnasmetleuaspoluser 584
Db 1721 tgcagagcgtgagcggcgccgagcgccgagcgccgagcgccgagcgccg 1755
QY 584 legllythrhispheaseranthrserleuvalthrserleusersearargly 604
Db 1756 --gcggcgacacatgagagcgagc-----acagacctgggacacatcc 1804
QY 604 laserprogluayarglypro-----serleuclue 615
Db 1805 agagccctgac--agggcgctgagcgcgagcgcgagcgcgagcgagcgag 1861
QY 615 heptrometproptrometgluthrlylevalasnprolllegllythrserval 635
Db 1862 tgcagagcggcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 1915
QY 635 rpleu--proserprothrvalgluclueargproserprola--lleaserleuser 653
Db 1916 gggccttgccttgccttgccttgccttgccttgccttgccttgccttgcct 1975
QY 653 lsleupovalphealasertrpasp 662
Db 1976 acctgccaatgttgcgcgagcgagcgagc 2004

```

```

RESULT 7
AY117207
LOCUS
DEFINITION
Arabidopsis thaliana putative ovule development protein
accession
AY117207
VERSION
AY117207.1
KEYWORDS
FLI CDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1699)
REFERENCE
1 Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, B., Dale, J.M.,
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
Tang, C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K.,
Davis, R.W., Ecker, J.R. and Theologis, A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 1699)
TITLE
JOURNAL
REFERENCE
AUTHORS
COMMENT
The Salk, Stanford, PGC (SSP) Consortium members constructed and
sequenced the pPNT (ORF) clones using the RAPL cDNAs: Yamada, K.,
Banh, J., Chan, M.M., Chang, C.H., Chang, B., Dale, J.M., Deng, J.M.,
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,
Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,
Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A.,
Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (03-JUN-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL CDNA: 'RIKEN
Arabidopsis Full-Length CDNA'): Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinzaki, K.
FEATURES
source
location/Qualifiers
1..1699
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U11873"
/ecotype="Columbia"
/note="This clone is in pPNT 51."
1..1699
/gene="At4g37750"
1..1668
/gene="At4g37750"
/codon_start=1
/evidence=experimental
/product="putative ovule development protein atnquegmenta"

```





Db 581 GGA----- 583  
 Qy 218 GINGlMetAsnCySglMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGly 237  
 Db 584 GGGCTTTTATATAGGGCTATATGGGAATTTCAA----- 619  
 Qy 238 CySglGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysVal 257  
 Db 620 -----CAGTCACAGCTTATCCATGAGCCCTGGGGCACAATCTAGCTGACATC 667  
 Qy 258 ThrAlaProSer----- 261  
 Db 668 ACTGGCTCTCACCAACCAACCAAAACCAAAACCAACCAACCAACCAACCAACCAAC 727  
 Qy 262 -----GlyThrAspSerValAlaVal 268  
 Db 728 CAGATCTCTGAGCTCTTGTGAGACACAGCGCTGGCTTGGACGACGACCAATGGCGGCT 787  
 Qy 269 AspaAlaIysIysArgGlyHis-----AlaIysLeuGlyGlnIysGlnProValHis 285  
 Db 788 GCGAAGAGAGAGAGAGAGAGAGATGTTGATGTTGGTACAGAAACAGATTGTTCAAT 847  
 Qy 286 ArgIysSerIleAspThrPheGlyGlnArgThrSerGlnIysArgGlyValThrArgHis 305  
 Db 848 AGAAATCTATCGATCTTTTGGACAGAACTTCTCAATACCGAGCGCTTACAGACAT 907  
 Qy 306 ArgGlyThrGlyArgGlyValAlaHisLeuThrAspAsnSerCysIysIysGlyGln 325  
 Db 908 AGATGGACTGGTATATGAAAGCTCATCTATGGACCAATGTTCAAGAAAGAGGTCTAC 967  
 Qy 325 ThrArgIysGlyArgGlnValIlyLeuGlyGlyIysAspMetGlnIysAlaAlaArg 345  
 Db 968 AGTGAAGAAAGAGACAGATTATCTGGAGCTTATATATGAGAGAAAGCTGCTCGCA 1027  
 Qy 346 AlaIysThrLeuAlaAlaLeuIysIysIysIysIysIysIysIysIysIysIysIys 365  
 Db 1028 GCATATGATCTTCTGCACTCACTACCTGCGGCTCCCTCTCTCTCAACCAATTTCTGCG 1087  
 Qy 366 GlnSerIysGlyGlnValGlnLeuGlnIysMetIysAsnMetSerArgGlnIysIysAla 385  
 Db 1088 GAGAAATATACAGAAAGATGAGACATGAGACATGAGACATGAGACATGAGACATGAG 1147  
 Qy 386 HisLeuArgArgIysSerSerGlyPheSerArgIysAlaSerIleIysArgGlyValThr 405  
 Db 1148 CATTGAG 1207  
 Qy 406 ArgHisIleGlnIleGlyArgIysIysIysIysIysIysIysIysIysIysIysIys 425  
 Db 1208 AGACATACCAAGCATGAG 1267  
 Qy 426 LeuIysIysLeuGlyIysThrPheSerThrGlnIysGlnIysAlaIysAlaIysAla 445  
 Db 1268 CTCATCTCTGGAACCTTTGGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1327  
 Qy 446 IleIysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgIysIysValGlu 465  
 Db 1328 ATTAAGTCCGTGGCAAAATGCTGTGATTAATCATCACAGAGTACAGAGATTGAT 1387  
 Qy 466 ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGlnLeuAlaArgGlyIysAspAsn 485  
 Db 1388 CGTATCATCTCTAGTACACACTCTGTCTGAGAGATTAGCGGAGAAC----- 1438  
 Qy 486 AspProArgAsnIysAspIleAspIysIysIysIysIysIysIysIysIysIysIys 505  
 Db 1439 -----AACCAAGCATTTCTC-----GTACAGGAATACT 1465  
 Qy 506 GlnThrValGlnValGlnAlaGlyAsnAsnAsnAsnGlnIysAspSerGlnIysMet 525  
 Db 1466 GAA----- 1468  
 Qy 526 ValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnIys 545  
 Db 1469 -----GACCAACACC 1477

Qy 546 IleMetAsnCySglMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGly 565  
 Db 1478 GCTCTAAATGCT----- 1489  
 Qy 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlnSerSerIleGly 585  
 Db 1489 ----- 1489  
 Qy 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGlyAlaSer 605  
 Db 1490 -----GTTGTGAGAGTGGTTCACCAAGAGTCACT 1522  
 Qy 606 ProGlnIysArgGlyProSerLeu-----LeuPheProMetProPheMetGlnThrIys 623  
 Db 1523 ACTCCCGAGAGACTCTTGAGTTTCCCGGCAATTTCCCGCTTCCCAAGTAAATCAAAAG 1582  
 Qy 624 IleVal--AsnProIleGlyThrSerValThrSerIleProSerProThrValGln 642  
 Db 1583 ATGTCGATCAATATGCGCGAATATGAGCTTGGACATCAACCACTTAATGCTGAG 1642  
 Qy 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTrpThrAsp 662  
 Db 1643 CTAAAG--ACCGTGGCTCTTACTTGTGCTCAATGCGGCTTTGCTGCTGGGCTGAT 1699  
 Qy 663 Thr 663  
 Db 1700 TCT 1702  
 RESULT 9  
 ATU04028  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana transcription factor CKC mRNA, complete cds.  
 ACCESSION  
 U44028  
 VERSION  
 U44028.1 GI:1171428  
 KEYWORDS  
 ORGANISM  
 Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 1955)  
 REFERENCE  
 Vergani, P., Morandini, P. and Soave, C.  
 Direct Submission  
 Submitted (29-DEC-1995) Paola Vergani, Dept. of Biology 'L.  
 Gorini', University of Milan, Via Celoria 26, Milan, 20133, Italy  
 location/Qualifiers  
 1..1955  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 76..1743  
 /function="transcription factor; complements a yeast  
 protein kinase C 1 mutant; Method: conceptual translation  
 supplied by author."  
 /codon\_start=1  
 /product="CKC"  
 /protein\_id="AA06281.1"  
 /db\_xref="GI:1171429"  
 /translation="MKSFCDDNDNNHSTNTLLGFSLSNMKKKGGRGREAIYSSST  
 CSAATSSSVLPOLVVGNTSNFVGCYSNPNPGI:YSMSVPLSDSLCMEALNR  
 SSHSNHODSPKEDPFGTHNNTSHKAMDLSDLSFYNTHEPNTTFOEPFSF  
 PQRHEETRYNDPGLTHGGSFNIVYVGFQSLSLMSPGSSCITGSHHQO  
 NONCHGSONHOIIEALVETSVGPEPTTMAAKKROEDVVVVGOKIYRKSIDT  
 FQCRSQRTGTRHWTRGKRYEHLMDNSFKRGSRKROYLLGGYDMEEKRAAYDL  
 AALKTWGSHTNFSNENYQKTEIMKMNTOEYLAHLRRKSSGSRASIRGCTVRL  
 HQHGWQALIGRVAENKDYLTFTFGQSEALAYVAIKRGINAATLRFDTTRDYD  
 RMSNTQLSGELARRNNNSIVRVTEQGLNNAVVEGSGNKEVSTPRLSLSPFIFA  
 LPQVNAQKFGSNMGMSPWTSNPNALKTVALTLDPQVFPVPAWMS"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6..9e-81 Length: 1955

Score: 1328.00  
 Percent Similarity: 52.84%  
 Best Local Similarity: 43.27%  
 Query Match: 37.92%  
 DB: 8  
 Matches: 312  
 Conservative: 69  
 Mismatches: 116  
 Indels: 224  
 Gaps: 21

US-10-024-632-2 (1-663) x ATU44028 (1-1955)

QY 1 MetLysArgIleAsnGluSerAsnAsnThrAspAspGlyAsnAsnHisAsnTrpLeuGly 20  
 DB 76 ATGAAGCTCTTTTGGATATGATGATATATATATATGCAACACAGACTAAATTGTGTAGG 135  
 QY 21 PheSerLeuSerProHis--MetLysMet----- 29  
 DB 136 TTCTCATTTGCTTCAAAATGATGATAAATGGAGGTAGAGAGGTAGAGAACTATTATAC 195  
 QY 30 -----GluAlaThrSerAlaAlaThrValProThrThrPheTyMet 43  
 DB 196 TCATCTTCACTTGTTCAGCTGCAACTCTTCTTCTGTTCACACTCAACTGTGTGT 255  
 QY 44 SerProSerGlnSerHisLeuSerAsnPheGlyMetCysTyArgValGlyGluAsnGly 63  
 DB 256 GGT-----GACACACTAGCAACTTGTGTGTCTGATGATCTAACCCAAATGGA 306  
 QY 64 AsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeuGlyLeu 83  
 DB 307 GGAATCTATCTTCACATCTCTGTGATGATGCACTCAGATCTGATGATCTCTTCTTATG 366  
 QY 84 GluAlaLeuLysArgSerGlnThrGlnValMetValProThrSerSerProLysLeuGlu 103  
 DB 367 GAAGCTCTCAACAGATCTCTCTCAGTCAGATCACCATCAAGATTATCTCCAAAGTGAG 426  
 QY 104 AspPheLeuGlyGlyAlaThrMetGlyThrHisGluTyArgSerHisGluArgGly--- 122  
 DB 427 GATTTCTTTGGG-----ACCATCAACAACAACAAGTCCAAAGAACCATG 474  
 QY 123 ---LeuSerLeuAspSerIleTyTrpAsnSerGlnAsnAlaGluAlaGlnProAsnArg 141  
 DB 475 GATCTTAGCTTAGATAGTTATTTATTCACAAACACTCAT-----GAGCCCAACAG 525  
 QY 142 AspLeuLeuSerGlnProPhe-----ArgGlnGlnGlyHisMetSerValGln 157  
 DB 526 ACTACAAACTTTCAGAGATCTTTAGCTTCCCTCAACCAACAACAT----- 573  
 QY 158 ThrHisProTyTrpSerGlyLeuAlaCysHisGlyLeuTyGlnAlaProLeuGlnGlu 177  
 DB 574 -----GAGGAA 579  
 QY 178 GluThrThrIleGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlu 197  
 DB 580 GAAACT---AGAAATTACGGGATGAC-----CTTAGTTTACACAT 618  
 QY 198 GlyLeuLysAsnTrpValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlu 217  
 DB 619 GGA----- 621  
 QY 218 GlnGlnMetAsnCysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGly 237  
 DB 622 GGGCTTTTAAATGAGGGGTATATGGGAATTTCAA----- 657  
 QY 238 CysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysVal 257  
 DB 658 -----CAATCACTAGCTTATTCATGACCTCTGGGTACAAATCTAGCTGATC 705  
 QY 258 ThrAlaProSer----- 261  
 DB 706 ACTGGCTCTCACACCACCAACAAACAAACAAACCAACCAAGCCAAACACACAG 765  
 QY 262 -----GlyThrAspSerValAlaVal 268  
 DB 766 CAGATCTCTGAAGCTCTTGTGAGACAAAGCGTTGGGTTTGAGACGACACAAATGGCGCT 825  
 QY 269 AspAlaLysLeuArgGlyHis-----AlaLysLeuGlyGlnLysGlnProValHis 285

DB 826 GCGAAGAAAGAAAGGGGACAAAGAGATTTGTAGTTGTGGTCAGAAACAGATTTTAT 885  
 QY 286 ArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyArgGlyValThrArgHis 305  
 DB 886 AGAAATCATGATGATCTTTGGACACAGAACTTCTCAATCCGAGGGCTTACAAAGACAT 945  
 QY 306 ArgTrpThrGlyValTyArgGluAlaHisIleuTrpAspAsnSerCysLysLysGlyGln 325  
 DB 946 AGATGAGCTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1005  
 QY 326 ThrArgLysGlyValArgGlnValTyArgGlyGlyTyArgPheMetGluGlyLysAlaArg 345  
 DB 1006 AGTAAAGAAAGAAAGCAAGTTTATGTGGAGTTATGATGAGAGAAAGCTCTGCA 1065  
 QY 346 AlaTyArgPheLeuAlaIleuLysTyTrpGlyProSerThrHisIleAsnPheSerIle 365  
 DB 1066 GCATATGATCTTGCAGCACTCAAGTACTGGGGTCCCTTACTCACACCAATTTCTGTGCG 1125  
 QY 366 GluAsnTyArgLysValGlnLeuGluGluMetLysAsnMetSerArgGlnGluTyValAla 385  
 DB 1126 GGAATTTACAGAAAGATTTAAGACATGAAAGACATGACTAGCAAGATATGTTGCA 1185  
 QY 386 HisLeuArgArgLysSerSerGlyPheSerArgLysAlaSerIleTyArgGlyValThr 405  
 DB 1186 CATTTGAGAAAGAAAGAGAGAGTGTCTTCTAGGGGTGCTTCATCTATATAGAGAGTACA 1245  
 QY 406 ArgHisIleGlnHisGlyValTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAsp 425  
 DB 1246 AGACATCACCAACATGAAAGGGTGGCAAGACAGATTTGATGATGCTGGAAACAAAGT 1305  
 QY 426 LeuTyArgGlyThrPheSerThrGlnGluGluAlaIleGluAlaTyArgValAlaAla 445  
 DB 1306 CTCTACCTTGAACCTTTGAAACCCAGAAAGAGCTGAGAAAGCTTACGATAGACAGA 1365  
 QY 446 IleLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyArgValGlu 465  
 DB 1366 ATTAAGTTCGGGACAAATGCTGTGACTTATTCATTCACAGAGGACATGATTTGAT 1425  
 QY 466 ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGluLeuAlaArgArgLysLysAspAsn 485  
 DB 1426 CGTATCATGCTAGTAAACACACTCTGTGTGAGAGTTAGCCGAAAGAAC----- 1476  
 QY 486 AspProArgAsnLysAspIleAspTyArgLysSerValValThrSerValAsnAsnGlu 505  
 DB 1477 -----AACCAACAGCATTTGTC-----GTCAGGAATACT 1503  
 QY 506 GluThrValGlnValGlnAlaGlyAsnAsnAsnGluAsnAspSerGluTrpLysMet 525  
 DB 1504 GAA----- 1506  
 QY 526 ValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLys 545  
 DB 1507 -----GACCAAAACC 1515  
 QY 546 IleMetAsnCysGlyAsnTyArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565  
 DB 1516 GCTCTAAAGCT----- 1527  
 QY 566 GlyIleAspSerValGlySerGlyHisAsnMetLeuAspGluSerSerLysIleGly 585  
 DB 1527 ----- 1527  
 QY 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGluLys 605  
 DB 1528 -----GTTTGGAGAGTGTTCACAAACAAAGATGACT 1560  
 QY 606 ProGluLysArgGlyProSerLeu-----LeuPheProMetProProMetGluThrLys 623  
 DB 1561 ACTCCGAGAGACTCTTGTGATTTCCGGGAGATTTTCGGGTGCTCAAGTTTATCAAAAG 1620  
 QY 624 IleVal---AsnProIleGlyThrSerValThrSerTrpLeuProSerProThrValGln 642

Db 1621 ATGTCGATCAATATGCGGAGAAATATGATGCTTGGACATCAACCTAATGCTGAG 1680  
 QY 643 MetarProserProalailSerleuserHieulProvalphealaserTPthraap 662  
 Db 1681 CTAAAG---ACCGGCTCTTACTTCTGCTCAAGATGCCGCTTTGCTGCTGCTGCTGAT 1737  
 QY 663 Thr 663  
 Db 1738 TCT 1740

RESULT 10  
 AY080706 2056 bp mRNA linear PLN 18-SSP-2002  
 LOCUS Arabidopsis thaliana putative ovule development protein  
 DEFINITION Aintegumenta (At4g37750) mRNA, complete cds.  
 ACCESSION AY080706 GI:19310586  
 VERSION AY080706.1  
 KEYWORDS Arabidopsis thaliana (chale crees)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana (chale crees)  
 Brachyotops; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 2056)  
 Yamaoka, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
 Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
 Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
 Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
 Saitou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K.,  
 Davis, R.W., Ecker, J.R. and Theologis, A.  
 Arabidopsis Full Length CDNA Clones  
 Unpublished  
 2 (bases 1 to 2056)  
 Yamaoka, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,  
 Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,  
 Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,  
 Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
 Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
 Sakurai, T., Saitou, M., Seki, M., Shim, P., Southwick, A.,  
 Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
 Direct Submission  
 Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan  
 Street, Albany, CA 94710, USA  
 RIKEN Genomic Sciences Center (GSC) members carried out the  
 collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN  
 Arabidopsis Full-length CDNA"): Seki, M., Narusaka, M., Ishida, J.,  
 Saitou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
 Hayashizaki, Y. and Shinzaki, K.

COMMENT  
 The Salik, Stanford, PGSC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAFL cDNAs: Yamaoka, K., Banh, J.,  
 Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,  
 Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
 Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chan, H.,  
 Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,  
 Meyer, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P.,  
 Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

FEATURES  
 source  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="4"  
 /clone="RAFL06-86-F19 (R11873)"

gene  
 5'UTR  
 misc\_difference  
 CDS  
 3'UTR  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7,39e-81 Length: 2056  
 Score: 1328.00 Matches: 312  
 Percent Similarity: 52.84% Conservative: 69  
 Best Local Similarity: 43.27% Mismatches: 116  
 Query Match: 37.92% Indels: 224  
 DB: 8 Gaps: 21

US-10-024-632-2 (1-663) x AY080706 (1-2056)  
 QY 1 MetlySAgileAeNGluserAsnAsnthzaspapglYaenAenHieAntryleugly 20  
 Db 158 ATGAAGCTCTTTTGGATGATGATGATATATATGATGCAACAGACATATTGTTGGG 217  
 QY 21 PheSerleuserProHis---MetlyeNet----- 29  
 Db 218 TTCTCATTTGCTTTCMAATATGATGAAATGGAGATGAGAGATGAGAACTATTAC 277  
 QY 30 -----GlualAthrSerAlaIathValProthThPhetyrmet 43  
 Db 278 TCATTTCACTTCTTCAAGTCGCAACTTCTTCTTCTTCTTCCACTCACTTGTGTT 337  
 QY 44 SerProserGlnSerHisleuserAsnBheGlymetCysYrGlyValGlyGluasnGly 63  
 Db 338 GGT-----GACAACTACAGTCAACTTGTGTGTGTCTATGATCTAACCAATGGA 388  
 QY 64 AenPheHisSerProleuThValMetProleuYsserAapglYserleuCytleleu 83  
 Db 389 GGATCTATTTCACATGCTGATGATGACACTCAATTTGATGATGTTCTTCTTCTTATG 448  
 QY 84 GlualAeulYargSerGlnThGlnValMetValProthSerSerProlyleuGlu 103  
 Db 449 GAAGCTTCAACAGATCTTTCATCGATCAACATCAAGATTCATCTCCAAAGGTGAG 508  
 QY 104 AspPheleuGlyValAlaThMetGlyThzHsGlyUryGlySerHisGluArggly 122  
 Db 509 GATTTCTTTGGG-----ACCATCAACAACAACAAGTCAACAAGACCATG 556  
 QY 123 ---LeuserleuAspSerIleTyTyZasnserGlnAenAlaGluAlaGlnProAsnArg 141  
 Db 557 GATCTTAGCTTAGATAGATTATATTCACAAACACACTCAT-----GAGGCCAACACG 607

/ecotype="Columbia"  
 /note="this clone is in a modified pluescript vector  
 (lambda ZAP) as a XhoI/SacI insert."  
 1..2056  
 /gene="At4g37750"  
 1..157  
 /gene="At4g37750"  
 2  
 /gene="At4g37750"  
 /note="not present in genomic sequence"  
 158..1825  
 /gene="At4g37750"  
 /codon\_start=1  
 /evidence=experimental  
 /product="putative ovule development protein Aintegumenta"  
 /protein\_id="At4g5024.1"  
 /db\_xref="GI:19310586"  
 /translation="MKSFCDDNDNNHNTTNLLGFSLSNNMKKGGREGREAITSSST  
 SSAATSSSVPOLVGDNTSNFVCYGSNPNNGIYSHMSVPLRSDGSLCMLALNR  
 SSSNHHQDSFPKVEDFGTHHTNTHSKEMADSLDLEFYTFTEPTNTTFOEFFR  
 POTRHEETRYNGNDPDLTHGSPFNQVYGEFQOGLSLSMSPGSSCITGSHHQ  
 NONHONHONHOQISALVETSVGETTTMAAAKKRGQOEIVVVGOKOIVHRSIDT  
 FGERTSOYRGYTRRPMRGVRAHLMDSFKKEGSRGROYIGGYDMEEAARAYDL  
 AALKRWGSPSTTNNSAERYQKEIEDKMTROQEVYALHRRSGSSGASLYRYVTTH  
 HQHRWQRIQIRVAGNDLYIGTGTEBAEAIVDAIKRGTNAVTNPFITRYVD  
 RIMSNLTLGSELARRNNNSIVYNTDQTLNANVVGASIKFVSTPERLTFPAIFA  
 LPQVOKMFGSLNMGNNMSPMSTNPAELKVALTLPPMPVFAAWADS"  
 1826..2056  
 /gene="At4g37750"



QY 142 AspleuteuserGlnPrope-----ArgGlnGlnGlyHisMetSerValGln 157  
 DB 608 ACTCAAAACCTTCAAGAGTCTTTAGCTTCCCTCAAAACGAAACCT----- 655  
 QY 158 ThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGln 177  
 DB 656 -----GAGGAA 661  
 QY 178 GluThrThrIysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGln 197  
 DB 662 GAAACT-----AAGAAATTACGGAAATGAC-----CTTACTGTGACCAT 700  
 QY 198 GlyLeuIysAsnTyrValAlaProThrArgGlnPheSerThrHisGlnGlnValLeuGln 217  
 DB 701 GGA----- 703  
 QY 218 GlnGlnMetAsnGlyMetGlyAsnGlnIuArgAsnGlyValSerLeuGlySerValGly 237  
 DB 704 GGGCTCTTTAATGTAGGGGTATATGGGAATTCA----- 739  
 QY 238 CysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysVal 257  
 DB 740 -----CAGTCACTGAGCTTATCCATGAGCCCTGGGTCAACATCTAGCTGCATC 767  
 QY 258 ThrAlaProSer----- 261  
 DB 788 ACTGGCTCTCACACCAACCAAAACCAAAACCAACCAACCAACCAACCAAC 847  
 QY 262 -----GlyThrAspSerValAlaVal 268  
 DB 848 CAGATCTCTAGAGCTCTTGAGACAACAGCTGGTTGAGACAGACAACAGAGGGCGCT 907  
 QY 269 AspalalysIysArgGlyHis-----AlaIysLeuGlyGlnIysGlnProValHis 285  
 DB 908 GCGAAGAAAGAGAGGGGACAGAGAGATGTGTAGTTGTGTGTGTGTGTGTGTGTGTGT 967  
 QY 286 ArgIysSerIleAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValThrArgHis 305  
 DB 968 AGAAATCTATCGATCTTTGGACACAGAACTTCTGATACCGAGCGCTTACAGACAT 1027  
 QY 306 ArgTyrThrGlyArgTyrGlyAlaHisIleuTyrAspAsnSerCysIysGlyGln 325  
 DB 1028 AGATGAGCTGCTAGATATGAACTCATCTATGAGACATAGTTTCAAGAGAGAGTGCAC 1087  
 QY 326 ThrArgIysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnIuValAlaArg 345  
 DB 1088 AGTGAAGAAAGAGACAGATTTATCTGGAGGTTATGATATGAGAGAAAGCTGCTGCA 1147  
 QY 346 AlaTyrAspLeuAlaIleuIleuTyrTyrProSerThrHisIleAsnPheSerIle 365  
 DB 1148 GCATATGATCTTGCGACACTCAAGTATCGGGGTCCCTCTCACTCACACCAATTTCTCGCG 1207  
 QY 366 GluMetTyrGlnValGlnLeuGlnGluMetIysAsnMetSerArgGlnGlnTyrValAla 385  
 DB 1208 GAGAAATTTACAGAAAGAGATTGAGACCTGAAAGAACATGATACAGAAATATGTTGCA 1267  
 QY 386 HisIleuArgArgIysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThr 405  
 DB 1268 CATTTGAGAGAGAGAGAGAGAGTGTCTTCTAGGGGTGCTTCATCTATGAGAGAGTGCACA 1327  
 QY 406 ArgHisHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnIysAsp 425  
 DB 1328 AGACATCCACAGCAGAGAGAGTGGACAGAGATGGTATAGTCTGGAGAAAGAT 1387  
 QY 426 LeuTyrLeuGlyThrPheSerThrGlnGlnIuAlaIleGlnAlaTyrAspValAlaAla 445  
 DB 1388 CTCTACCTTGGACCTTTTGAACCCAGAGAGAGAGAGAGCTTCAGATGTAGCGACA 1447  
 QY 446 IleIysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrAspValGlu 465  
 DB 1448 ATTAAGTTCGGTGCAAAATGCTGTGCTAATCTTGTATATACAGAGGTAGAGATGTTGAT 1507  
 QY 466 ArgIleMetAlaSerSerAsnIleuValGlyGluLeuAlaArgAlaGlyIysAspAsn 485

DB 1508 CGTATCATGTCTAGTAACACACTCTTGCTCGAAGATTAAGCGCAAGAAC----- 1558  
 QY 466 AspProArgAsnIysAspIleAspTyrAsnIysSerValValThrSerValAlaAsnGlu 505  
 DB 1559 -----AACACAGCATTTGTC-----GTCAAGAAATACT 1585  
 QY 506 GluThrValGlnValGlnAlaGlyAsnAsnAsnAsnGlnAsnAspSerGluTyrIysMet 525  
 DB 1586 GAA----- 1588  
 QY 526 ValIleuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnIys 545  
 DB 1589 -----GACCAAAAC 1597  
 QY 546 IleMetAsnGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565  
 DB 1598 GCTCTAATGCT----- 1609  
 QY 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerSerIleGly 585  
 DB 1609 ----- 1609  
 QY 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgIuAlaSer 605  
 DB 1610 -----CTTGTGAGAGTGTTCACAAAGAACTCAGT 1642  
 QY 606 ProGluIysArgGlyProSerLeu-----LeuPheProMetProProMetGluThrIys 623  
 DB 1643 ACTCCGAGAGACTTGTAGTTTCCGCGAGATTTCCGGTTCCTCAAGTATATCAAAAG 1702  
 QY 624 IleVal-----AsnProIleGlyThrSerValThrSerTyrLeuProSerProThrValGln 642  
 DB 1703 ATGTTGGATCAAAATATGAGGAGAAATGAGTCTTGAATCAAAACCTTAAGCTGAG 1762  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisIleuProValPheAlaSerTyrThrAsp 662  
 DB 1763 CTTAAG---ACCGTGTCTTACTTGTGCTCAGATGCCGTTTGTGCTGCTGGAGT 1819  
 QY 663 Thr 663  
 DB 1820 TCT 1822  
 RESULT 11  
 BD274516 2148 bp DNA linear PAT 17-JUN-2003  
 LOCUS BD274516  
 DEFINITION Methode for modifying weight of organ, controlling fertility and enhancing asexual reproduction in plant.  
 ACCESSION BD274516  
 VERSION BD274516.1 GI:33084284  
 KEYWORDS JP 2002534078-A/1.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 2148)  
 Fischer,R.L. and Mizukami,Y.  
 Methode for modifying weight of organ, controlling fertility and enhancing asexual reproduction in plant  
 Patent: JP 2002534078-A 1 15-OCT-2002;  
 THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
 Arabidopsis thaliana (thale cress)  
 OS Arabidopsis thaliana (thale cress)  
 PN JP 2002534078-A/1  
 PD 15-OCT-2002  
 PF 07-JAN-2000 JP 2000592392  
 PR 08-JAN-1999 US 09/227421  
 PI ROBERT L FISCHER, YUKIKO MIZUKAMI  
 PC A01H1/00, C12N5/10, C12N15/09, C12N5/00, C12N15/00 CC  
 AINTEGUMENTA (AMT)  
 CC encodes first AP2 domain  
 CC encodes linker region  
 CC encodes second AP2 domain





QY 566 GlyIleAspSerValGlySerGlyGlnHisAsnValLeuAspGlnUserSerLysIleGly 585  
 Db 1720 ----- 1720  
 QY 586 ThrHisPheSerAsnThrSerSerSerLeuValThrSerLeuSerSerAsnGluAlaSer 605  
 Db 1721 -----GTTGTGGAAGGTGGTTCACCAAGAAAGTCACT 1753  
 QY 606 ProGluYAspArgGlyProSerLeu-----LeuPheProMetProMetGluThrLys 623  
 Db 1754 ACTCCCGAGAGACTCTTGAAGTTTCGCCGCAATTTTCGGTTCCTCAAGTAATCAAAAG 1813  
 QY 624 ILeuVal-----AsnProIleGlyThrSerValThrSerThrLeuProSerProThrValGln 642  
 Db 1814 ATGTCGGATCAAAATATGCGGAAATATGAGTCCTTGACATCAAAACCTTAATGCTGAG 1873  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerThrThrAsp 662  
 Db 1874 CTTAAG-----ACCGTCGCTCTTACTTTCCTCAAGATCCGGTTTCGCTGCTGGAGCTGAT 1930  
 QY 663 Thr 663  
 Db 1931 TCT 1933  
 RESULT 12  
 AR316367 2148 bp DNA linear PAT 12-JUN-2003  
 LOCUS AR316367  
 DEFINITION Sequence 1 from patent US 6559357.  
 ACCESSION AR316367  
 VERSION AR316367.1 GI:31711158  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 2148)  
 AUTHORS Fischer, R.L. and Mizukami, Y.  
 TITLE Methods for altering mass and fertility in plants  
 JOURNAL Patent: US 6559357-A 1 06-MAY-2003;  
 FEATURES  
 source location/Qualifiers  
 1..2148  
 /organism="unknown"  
 /mol\_type="genomic DNA"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7,846-81 Length: 2148  
 Score: 1328.00 Matches: 312  
 Percent Similarity: 52.84% Conservative: 69  
 Best Local Similarity: 43.27% Mismatches: 116  
 Query Match: 37.92% Indels: 224  
 Gaps: 21  
 Db: 6  
 US-10-024-632-2 (1-663) x AR316367 (1-2148)  
 QY 1 MetLysArgGluAsnGlnSerAsnAsnThrAspArgGlyAsnAsnHisAsnTrpLeuGly 20  
 Db 269 ATGAAGTCTTTTGTGATATGATGATATATATATATATATATATATATATATATATATAT 328  
 QY 21 PheSerLeuSerProHis-----MetLysMet----- 29  
 Db 329 TTCTCATTTCTTCAATATGATGATAAATGGAGGTAGAGAGAGTAGAAGAGCTAATTAC 368  
 QY 30 -----GluAlaThrSerAlaAlaThrValProThrThrPheThrMet 43  
 Db 389 TCATCTTCAACTTCTTCACTGCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 448  
 QY 44 SerProSerGlnSerHisLeuSerAsnPheGlyMetCysTyrGlyValGlyLysAsnGly 63  
 Db 449 GGT-----GACAACTAGCAACTTGTGTGCTGCTAGTATGATCAACCAAAATGA 499  
 QY 64 AsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeuCysIleLeu 83  
 Db 500 GGAATCTATCTCACTATCTGTGTGATGCGCACTCAATCTGATGATGATGATGATGATGATG 559

QY 84 GluAlaLeuLysArgSerGlnThrGlnValMetValProThrSerSerProLysLeuGlu 103  
 Db 560 GAAGCTCTCAACAGATCTTCTCAGTCGATACCACTCAAGATTCATCTCCAAAGGTGAG 619  
 QY 104 AspPheLeuGlyValAlaThrMetGlyThrHisGlyTyrGlySerHisGluArgGly 122  
 Db 620 GATTCCTTTGGG-----ACCATCAACAACAACAGTCACAAAGAACCCATG 667  
 QY 123 ---LeuSerLeuAspSerIleTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsnArg 141  
 Db 668 GATCTTAGCTTAGATAGATTATCTCAACACCACTACT-----GAGCCCAACAG 718  
 QY 142 AspLeuLeuSerGlnProPhe-----ArgGlnGlnGlyHisMetSerValGln 157  
 Db 719 ACTCAAACTTCAAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 766  
 QY 158 ThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGlu 177  
 Db 767 -----GAGGA 772  
 QY 178 GluThrThrLysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlu 197  
 Db 773 GAAACT-----AGAAATACGGGAATGAC-----CTTAAGTTTGCACACT 811  
 QY 198 GlyLeuLysAsnTrpValAlaProThrThrArgLysPheSerThrHisGlnGlnValLeuGlu 217  
 Db 812 GGA----- 814  
 QY 218 GlnGlnMetAsnCysGlyMetGlyAsnGlnArgAsnGlyValSerLeuGlySerValGly 237  
 Db 815 GGGCTTTTAATGATGAGGGTATATGCGAATTCAA----- 850  
 QY 238 CysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysVal 257  
 Db 851 -----CAGCTAGCTGAGCTTATCCAGAGCCCTGGGTCAAACTTAAGCTGACATC 898  
 QY 258 ThrAlaProSer----- 261  
 Db 899 ACTGCGCTCACACACACACACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 958  
 QY 262 -----GlyThrAspSerValAlaVal 268  
 Db 959 CAGATCTGAAGCTCTTGTGAGACAAAGCGTGGCTTTGAGACAGCAACATGCGGCT 1018  
 QY 269 AspAlaLysLysArgGlyHis-----AlaLysLeuGlyGlnLysGlnProValHis 285  
 Db 1019 GCGAAGAAAGAAAGGAGGACAAAGAGATGTGTAGTTGTGTGTGTGTGTGTGTGTGTGTGT 1078  
 QY 286 ArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValAlaThrArgHis 305  
 Db 1079 AGAAATCTATGATGATCTTTTGACAAACGAACTTCTCAATACCGAGCGCTTCAAGACAT 1138  
 QY 306 ArgTrpThrGlyArgTyrGlnAlaHisLeuTrpAspAsnSerCysLysLysGlnGluGln 325  
 Db 1139 AGATGACCTGGGAGATATAGATCACTCATCTATGAGCAAAAGTTTCAAGAAAGAGCTCAC 1198  
 QY 326 ThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnGluLysAlaAlaArg 345  
 Db 1199 AGTAAAGAAAGAAAGAAAGAAAGTATATCTGAGAGCTTATGATATGAGAGAGAAAGCTGTCCA 1258  
 QY 346 AlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyProSerThrHisIleLeuPheSerIle 365  
 Db 1259 GCAATATGATCTTGTGCTCACTCAAGTACTGGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1318  
 QY 366 GluAsnTyrGlnValGlnLeuGlnGluLysMetLysAsnMetSerArgGlnGlnTyrValAla 385  
 Db 1319 GAGAAATTAACAGAAAGAGATTGAGACATGAGAAACATGACTAGACAGAAATATATGTTGGA 1378  
 QY 386 HisLeuLysArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThr 405  
 Db 1379 CATTGAGAAAGAAAG 1438



QY 262 -----GlyThrAspSerValAlaVal 268  
 Db 959 CAGATCTCTGAAGCTCTTGGAGACAGCGCTTGGAGCGCGCAATGGCGGCT 1018  
 QY 269 AspaIaLysLysArgGlyHis-----AlaLysLeuGlyGlnLysGlnProValHis 285  
 Db 1019 GCGAAGAGAGAGAGGAGAGAGATGTGTGATGTTGCTGCTCAGAAAACAGATGTTCT 1078  
 QY 286 ArgLysSerLysLeuArgPheGlyGlnArgThrSerGlnThrArgGlyValThrArgHis 305  
 Db 1079 AGAAATCTATCGATCTTGGACACGAACTTCAACCGAGCGCTTCAAGACAT 1138  
 QY 306 ArgTTPheGlyArgTyrGlnAlaHisLeuTTPheAspAsnSerCysLysLysGlnGln 325  
 Db 1139 AGATGAGCTGGTATATGAAAGCTCAATCTATGGACAAATGTTTCAAGAGAGAGCTCAC 1198  
 QY 326 ThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnGlnLysAlaAlaArg 345  
 Db 1199 AGTGAAGAGAGAGACAGATTATCTGGGAGGTTAGATAGAGAGAGAACTCTCTGA 1258  
 QY 346 AlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyProSerThrHisLeuAspSerIle 365  
 Db 1259 GCATATGATCTTGGCTGCACTCAAGTACTGGGGTCCCTTACACACCAATTTCTCTGG 1318  
 QY 366 GlnAsnTyrGlnValGlnLeuGlnMetLysAspMetSerArgGlnGlnLysValAla 385  
 Db 1319 GAGATATATCAGAAAGAGATTGAGACATGAGACATGACATGACAAATATTTCTCA 1378  
 QY 386 HisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThr 405  
 Db 1379 CATTTGAGAGAGAGACAGTGGTTTCTTCAAGGGGCTTCTCATCTATGAGAGAGTACA 1438  
 QY 406 ArgHisHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlnLysLysAsp 425  
 Db 1439 AGACATACACAGCATGGAGAGTGCACAGCAGATGTGAGAGTGCCTGAGAACAAAT 1498  
 QY 426 LeuTyrLeuGlyThrPheSerThrGlnGlnGlnAlaAlaGlnAlaTyrAspValAlaAla 445  
 Db 1499 CTACCTCTGGAACCTTTGACACCAAGAGAGAGCTGACAGAGCTTACATGATACACA 1558  
 QY 446 IleLysPheArgGlyAlaAsnAlaValThrAspPheAspIleSerArgTyrAspValAlu 465  
 Db 1559 ATTAGATTCCTGGCACAAATGCTGCTGACTTAACTTGAATACCGAGAGCATGTGAT 1618  
 QY 466 ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGlnLeuAlaArgArgLysLysAspAsn 485  
 Db 1619 CGATCATGCTCTAGTACACACTCTTGTCTGAGAGATGACCGAAGAAC----- 1659  
 QY 486 AspProArgAsnLysAspIleLeuTyrAsnLysSerValValThrSerValAsnAsnGln 505  
 Db 1670 -----MACACACGATTTCTC-----GTCAGCAATTTCT 1696  
 QY 506 GlnThrValGlnValGlnAlaGlyAsnAsnAsnAsnGlnLysAspSerGlnTyrLysMet 525  
 Db 1697 GAA----- 1699  
 QY 526 ValLeuPheAsnHisPheProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLys 545  
 Db 1700 -----GACCAAAACC 1708  
 QY 546 IleMetAsnCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565  
 Db 1709 GCTCTAAATGCT----- 1720  
 QY 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlnSerSerLysIleGly 585  
 Db 1720 ----- 1720  
 QY 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerArgGlnAlaSer 605  
 Db 1721 -----GTTGTGGAAGGTGTTTCCAAACAAAGAAAGTCACT 1753

QY 606 ProGlnLysArgGlyProSerLeu-----LeuPheProMetProProMetGlnThrLys 623  
 Db 1754 ACTCCCGAGAGACTCTTGAGTTTCCGCGGATTTTCGCGTCTTCAAGATTAATCAAAAG 1813  
 QY 624 IleVal---AsnProIleGlyThrSerValThrSerTyrLeuProSerProThrValGln 642  
 Db 1814 ATGTCGATCAATCAATGTGGCGGAATATGATGATCTTGCATCAACCTTAAGCTGAG 1873  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTyrThrAsp 662  
 Db 1874 CTTAAG---ACCGTCTCTTACTTGTGCTCAAGATCGGATTTGCTGCTGGCTGAT 1930  
 QY 663 Thr 663  
 Db 1931 TCT 1933  
 RESULT 14  
 ATU40256 2148 bp mRNA linear PLN 29-FEB-1996  
 LOCUS Arabidopsis thaliana ovule development protein (AINTEGUMENTA) mRNA,  
 DEFINITION complete cds.  
 ACCESSION U40256.1 GI:1209098  
 VERSION U40256  
 KEYWORDS  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 2148)  
 REFERENCE Klucher,K.J., Chow,H., Reiser,L. and Fischer,R.J.  
 AUTHORS The AINTEGUMENTA Gene of Arabidopsis is Required for Ovule and Female  
 TITLE Gametophyte Development is Related to the Floral Homeotic Gene  
 APTPLA2  
 JOURNAL Plant Cell 8 (1996) In press  
 REFERENCE 2 (bases 1 to 2148)  
 AUTHORS Klucher,K.  
 TITLE Direct Submision  
 JOURNAL Submitted (07-NOV-1995) L. Reiser, Plant Biology, UC Berkeley, 111  
 Koshland Hall, Berkeley, CA 94720, USA  
 FEATURES  
 source  
 1..2148  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="4"  
 1..2148  
 /gene="AINTEGUMENTA"  
 259..1936  
 /gene="AINTEGUMENTA"  
 /note="AP2 domain containing protein"  
 /codon\_start=1  
 /product="AINTEGUMENTA"  
 /protein\_id="AAA91040.1"  
 /db\_xref="gi:1209098"  
 /translation="MKSFDNDNDNNHNTNLLGPSLSSNMKKMGKGGGREGRAIVSSST  
 SGAATSSVLPQLVVGDNTSGFVCGVSGNSNGGIYSMSMPLRSGVSLCLIMALNR  
 SSSNNHDDSPKVEDFRCGTHNNNTSHKENDSLDI.FVNTTEPNTTNPOFFSF  
 PQRNHEERTRYGNDPSLTGSGFNNGVYEFQCSLSLSPSSQSCITGSHHHC  
 NONNHHGSHQQLSEALVETSVGFTTMAAKKRGQEDVVVGQQLVHRKSLDT  
 FQORTQIRGVTRHWTRGVBAHLMNDSEFKEGSRKRGQVYLLGGMESGAARAYDL  
 ALKKTYSQIRHTNPSAENYQKEIIDMKMTQGEVVAHLRRKSGFSRGSATYRGVTR  
 HQHGRWQARIGVAGNKLVLGTGTQEEAAEDVAIVAKRGTAATVNTNPDITREYDV  
 RIMSSNTLLSGELARRNNNS.IVNTNTEQTLNVAIVGGSVKEVSTPERLLSPALPA  
 LFOVNGKRGFSNMGMGMSPWTSNPAELKTVALTLPQMPVFAAADS"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.84e-81 Length: 2148  
 Score: 1328.00 Matches: 312  
 Percent Similarity: 52.84% Conservative: 69  
 Best Local Similarity: 43.27% Mismatches: 116  
 Query Match: 37.92% Indels: 224

DB:	8	Gaps:	21
US-10-024-632-2 (1-663) x ATU40256 (1-2148)			
QY	1	MellyeRgllleasncluserbenanthrAapapglYaenbanhI:saentPleuGly	20
Db	269	ATGAAGCTTTTGTGTATATGATGATATATATCATTCAGCAACAGACTAATTTGTTAGGG	328
QY	21	PheSerLeuSerProHis--MellysMet	29
Db	329	TTTCATTTGCTCCAAATATGATGAAAGGAGAGGTAGAGAGGTATGACATTTTAC	368
QY	30	-----GluAlaThrSerAlaAlaAlaThrValProThrThrPheTyMet	43
Db	389	TCATCTTCAACTTCTTCAGCTGCACACTTTCTTTCTGTGTCCACTCACTGTGTT	448
QY	44	SerProSerGlnSerLeuSerAenPheGlyMetCySTyTgLYValGlyGluangly	63
Db	449	GGT-----GACAACTAGCAACTTTGGTGTTCATGATCTAACCCAAATGGA	499
QY	64	AsnPheHisSerProLeuThrValMetProLeuLysSerAapglYserLeuCySleleu	83
Db	500	GGAATCTATTTCTCAACTGCTGTGGAAGCCACATCAAGATGTGATGTTCTTCTCTTAAG	559
QY	84	GluAlaLeuLysAerGserGlnThrGlnValMetValProThrSerSerProLysLeuGlu	103
Db	560	GAACCTCTCAACAGACTTCTCTCACTCGAATCACCATCAAGATTCATCTCCAAAGGTGAG	619
QY	104	AspPheLeuGlyGlyAlaThrMetGlyThrHISgluTyrgLYserHISgluargGly--	122
Db	620	GATTTCTTTGGG-----ACCATCTCAACAAACACAGTCCAAAGAAAGCCATG	667
QY	123	---LeuSerLeuAspSerIleTyTyrAenSerGlnAsnAlaGluAlaGlnProAsnArg	141
Db	668	GATCTTAGCTTAGATGATTATTATCTCAACACACATCTAT-----GAGCCCAACAG	718
QY	142	AspLeuLeuSerGlnProPhe-----ArgGlnGlnGlyHISMetSerValGln	157
Db	719	ACTACAAACTTCAAGAGTTCTTTAGCTTCCCTCAACCAAGAAACCAT-----	766
QY	158	ThrHisProTyTyrSerGlyLeuAlaCysHisGlyLeuTyrgInAlaProLeuGlnGlu	177
Db	767	-----GAGGAA	772
QY	178	GluThrThrLysGlnThrHisValSerAapCySerSerLeuMetProGlnMetThrGlu	197
Db	773	GAACCT-----AGAAATTACGGGAATGAC-----CCTAGTTGACCAT	811
QY	198	GlyLeuLysAenTyrValAlaProThrArgGluPheSerThrHISGlnGlnValLeuGlu	217
Db	812	GGA-----	814
QY	218	GlnGlnMetAenCyGlyMetGlyAenGluArgAenGlyValSerLeuGlySerValGly	237
Db	815	GGCCTCTTTAATGAGGCGATATGGGGAATTCCA-----	850
QY	238	CysGlyGlnLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCyVal	257
Db	851	-----CAGTCACTGAGCTTATCCATGAGACCTCGGTCACAATCTTACCTGATC	898
QY	258	ThrAlaProSer-----	261
Db	899	ACTGGCTCTACACACACCAACAAACAAACAAACAAACCAACAAAGCCACACAG	958
QY	262	-----GlyThrAapSerValAlaVal	268
Db	959	CAGATCTCTGAAGCTCTTGTGAGACCAAGCGTTGGGTTTGAAGACGACCAATGGCGCT	1011
QY	269	AspAlaLysLysAerGlyHis-----AlaLysLeuGlyGlnTyrgInProValHis	285
Db	1019	GCAGAAAGAAAGAGGGGACAAAGAGATGTTGATGTTGGTCAGAAACAGATGTTGATC	1071
QY	286	ArgTySerIleAenThrPheGlyGlnArgThrSerGlnTyrgArgGlyValThrArgHis	305

Db	1079	AGAAATCTATCGATACCTTTGGACCAAGAACTTCTGAAATACGAGGCGTTACAAGCAT	1138
QY	306	ArgTrpThrGlyValGlyTrpGluAlaHisLeuThrPheAspAsnSerCysValGlyGluGlyGln	325
Db	1139	AGATGGACGTGTATATGAAAGCTCACTCTATGGACAAATAGTTTCAAGAAAGAAAGCTAC	1198
QY	326	ThrArgGlySerGlyArgGlnValTrpLeuGlyGlyGlyTrpAspMetGluIuValAlaAlaArg	345
Db	1199	AGTGAAGAAAGGAAGACAAGATTATCTCGAGAGGTTATATATGAGAGGAAAGCTGTGCA	1258
QY	346	AlaArgPheLeuMetAlaLeuValTrpTrpGlyProSerThrHisLeuPheSerIle	365
Db	1259	GCAATATGATCTGTGCGACCTCAATATCTAGGGGTCCCTCTACTCACACCAATTTCTGTGG	1318
QY	366	GluAsnTrpGlnValGlnLeuGluGluMetLeuAsnMetSerArgGlnGluTrpValAla	385
Db	1319	GAGAAATTATCGAAGAAAGATGTAAGAAACATGACTAGACAAGAAATATGTTGCA	1378
QY	386	HisLeuAsnArgValSerSerGlyPheSerArgGlyAlaSerIleTrpArgGlyValThr	405
Db	1379	CATTGKAGAGAAAGACAGTGGTTCTTCAAGGTCTTCCATCTTCAAGAGGTACA	1438
QY	406	ArgHisHisGlnHisGlyArgTrpGlnAlaArgIleGlyValAlaGlyAsnIleAsp	425
Db	1439	AGACATCACCCGATCGAGAGGTGGCAAGCACGCGATTGTAGAGTGGCTGGAAACAAAGAT	1498
QY	426	LeuTrpLeuGlyTrpPheSerThrGlnGlnGluAlaGluAlaTrpAspValAlaAla	445
Db	1499	CTTACCTTGGAACTTTTGGAAACCCAGAAAGAGCTCAGAGCTTTCGATGTAAGACCA	1558
QY	446	IleLeuPheArgGlyAlaAlaAlaValThrAsnPheAspIleSerArgTrpAspValGlu	465
Db	1559	ATTATAGTTCGGTGGCAACAAATGCTGTGATCAACTTATATATCAGAGTTCAGATGTGAT	1618
QY	466	ArgIleMetAlaSerSerAsnLeuLeuAlaGlyValuLeuAlaArgArgLysAspAsn	485
Db	1619	CGTATCATGTCTAGTAAACACACTTGTCTCGAGAGTTAGCCGCAAGAAC	1669
QY	486	AspProArgAsnIleAspIleAspTrpAsnLysSerValAlaThrSerValAsnAsnGlu	505
Db	1670	-----AACAAACGCACTTGTGC-----GTCCAGAAATACT	1696
QY	506	GluThrValGlnValGlnAlaGlyAsnAsnAsnAsnGluAsnAspSerGluTrpLysMet	525
Db	1697	GAA-----	1699
QY	526	ValLeuPheAsnHisProSerGlnGlnGlnGlnAlaAsnGluAsnGlySerAspGlnLys	545
Db	1700	-----GACCAAAACC	1708
QY	546	IleMetAsnGlySerGlyAsnTrpArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle	565
Db	1709	GCTCTAATGCT	1720
QY	566	GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerLysIleGly	585
Db	1720	-----	1720
QY	586	ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerArgGluAlaSer	605
Db	1721	-----GTTGTGGAAGTGTCTCAACAAAGAAAGTCAGT	1753
QY	606	ProGluLysArgGlyProSerLeu-----LeuPheProMetProProMetGluThrLys	623
Db	1754	ACTCCCGAGACACTTGAAGTTTCCGCGCAATTTCCGCTGTCCCTCAAGTAAATCAAAAG	1813
QY	624	IleVal-----AsnProIleGlyThrSerValThrSerTrpLeuProSerProThrValGln	642
Db	1814	ATGTTCCGATCAAAATATGCGCGGAAATATAGTCTTGCAATCAAAACCCATATGCTGAG	1873
QY	643	MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTrpThrAsp	662



```
Db 1192 GCTGGAACAAAGATCTCTACCTTGGAACCTTCGAACTCAAGAAAGCGCGGAGCC 1251
QY 441 TyrAspValAlaAlaIleIysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSer 460
Db 1252 TATGATGTAGCAGCTATCAAGTTCCGTGGCACAAACGCGGTGACTAATTGACATAACA 1311
QY 461 ArgTyrAspValGluArgGlyIleMetAlaSerSerAsnLeuValGlyGluLeuAlaArg 480
Db 1312 AGGTACAGATGTTATGCTAATAGCTAATGCTAATACACTCTTGTCTGGAAGATGCTCGA 1371
QY 481 ArgIleLysAspAsnAspProArgAsnLysAspIleAspTyrAsnLysSerValValThr 500
Db 1372 AGGAAACGCGC-----AACGCACTCGTGTCCGC 1398
QY 501 SerValAsnAsnGlnGlu-----ThrValGlnValGlnAlaGlyAsnAsnAsnAsn 517
Db 1399 AACATTAGCGACGAGAGCGCGCTTAAACCGCTGTCTGTAACGCTGTTCCAAATAAGGA 1458
QY 518 GluAsnAspSerGluTrpLysMetValLeuPheAsnHisProSerGlnGlnGlnAla 537
Db 1459 GTG----- 1461
QY 538 AsnGlyAsnGlySerAspGlnLysIleMetAsnGlyAsnTyrArgAsnSerAlaPhe 557
Db 1462 -----GGTAGCCCGGAGAGGTTTGAGTTT----- 1488
QY 558 SerMetAlaLeuGlnAspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMet 577
Db 1488 ----- 1488
QY 578 LeuAspGluSerSerLysIleGlyThrHisPheSerAsnThrSerSerLeuValThrSer 597
Db 1488 ----- 1488
QY 598 LeuSerSerSerArgGluAlaSerProGluLysArgGlyProSerLeuLeuPheProMet 617
Db 1489 -----CCGACGATATTGCGTTG 1506
QY 618 ProPheMetGluThrLysIle-----ValAsnProIleGlyThrSerValThrSerTrp 635
Db 1507 CCTCAAGTTGGTCCGAGAGATGTCGAGCAAAATGTGTCCGA--AATATGAGTTCTTGG 1563
QY 636 LeuProSerProThrValGlnMetArgProSerProAlaIleSerLeuSerHisLeuPro 655
Db 1564 ACTACGAACCTTAATGCTGATCTCAAG--ACCGTTCTTACTCTGCCGCAATGCCG 1620
QY 656 ValPheAlaSerTrpThrAspThr 663
Db 1621 GTTTTCGCTGCGTGGCTGATTTCT 1644
```

Search completed: March 13, 2004, 07:10:13  
Job time : 4737 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model.

Run on: March 13, 2004, 07:10:41 ; Search time 444 Seconds (with about 31 segments)

5497.521 Million cell updates/sec

Title: US-10-024-632-2  
Perfect score: 3502  
Sequence: 1 MKRINESNNTDDGNHHNWLG.....RSPAISLSHLVFASWTD 663

Scoring table:	
BLOSUM62	
Xgapop 10.0 ,	Xgapext 0.5
Ygapop 10.0 ,	Ygapext 0.5
Fgapop 6.0 ,	Fgapext 7.0
Delop 6.0 ,	Delext 7.0

Searched: 2432557 segs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

```

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/gen2.1/USFPO.spool/US10024632/runat.09032004.101210.10885/app.query.fasta.1.839
-DB=publishedApplications NA -CPWT=fasterp -SUFFIX=rmpb -MINMATCH=0.1
-LOOCC1=0 -IOOCC2=0 -UNITS=bits -START=1 -END=1 -MATRIX=blscom62
-TRANS=human40.cdd -LIST=45 -DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=5 -MODE=IOCALL -OUTPMT=ptc -NORM=ext -HEATSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10024632 @CGN.1.1.164 @unat.09032004.101210.10885
-NCPU=6 -ICPU=3 -NO MAP -LARGEOUTER -NEG_SCORES=0 -WAIT -DSPBIOCK=100
-LONGIO=6 -DEV.TIMOUT=120 -WARN TIMEOUT=30 -THRSAD=0 -XGAPEX=0.5
-XGAPEX=6 -FGAPEX=7 -YGAPEX=10 -YGAPEXT=0.5 -DEL0=6 -DELXT=7

```

Database : Published Applications NA:\*

1. /cgn2\_6/prodataa/2/pubnpna/US07\_PUBCOMB.seq.\*
2. /cgn2\_6/prodataa/2/pubnpna/PCr\_NEM\_PUB.seq.\*
3. /cgn2\_6/prodataa/2/pubnpna/US06\_NEM\_PUB.seq.\*\*
4. /cgn2\_6/prodataa/2/pubnpna/US06\_PUBCOMB.seq.\*\*
5. /cgn2\_6/prodataa/2/pubnpna/US07\_NEM\_PUB.seq.\*
6. /cgn2\_6/prodataa/2/pubnpna/PCrUS\_PUBCOMB.seq.
7. /cgn2\_6/prodataa/2/pubnpna/US08\_NEM\_PUB.seq.\*
8. /cgn2\_6/prodataa/2/pubnpna/US08\_PUBCOMB.seq.\*\*
9. /cgn2\_6/prodataa/2/pubnpna/US09a\_PUBCOMB.seq.
10. /cgn2\_6/prodataa/2/pubnpna/US09c\_PUBCOMB.seq.
11. /cgn2\_6/prodataa/2/pubnpna/US09c\_PUBCOMB.seq.
12. /cgn2\_6/prodataa/2/pubnpna/US09\_NEM\_PUB.seq.\*
13. /cgn2\_6/prodataa/2/pubnpna/US10a\_PUBCOMB.seq.
14. /cgn2\_6/prodataa/2/pubnpna/US10c\_PUBCOMB.seq.
15. /cgn2\_6/prodataa/2/pubnpna/US10\_NEM\_PUB.seq.\*
16. /cgn2\_6/prodataa/2/pubnpna/US00\_NEM\_PUB.seq.
17. /cgn2\_6/prodataa/2/pubnpna/US00\_PUBCOMB.seq.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
---------------	-------	----------------	--------	----	----	-------------

1	3502	105.0	2344	13	US-10-024-632-1	Sequence 1, Appl
2	2193.5	62.6	2405	12	US-10-424-139-135491	Sequence 136491, A
3	1744	49.8	1378	12	US-10-425-114-18711	Sequence 18711, A
4	1715	49.0	2323	13	US-10-024-632-2	Sequence 3, Appl1
5	1408.5	40.2	1926	13	US-10-024-332-5	Sequence 5, Appl1
6	1349.5	38.5	2010	13	US-10-024-332-8	Sequence 8, Appl1
7	1338	37.9	2148	14	US-10-059-911-1	Sequence 1, Appl1
8	1332	37.7	1668	14	US-10-059-911-39	Sequence 39, Appl1
9	1319	37.3	1668	14	US-10-059-911-38	Sequence 38, Appl1
10	1307	37.3	1359	14	US-10-059-911-4	Sequence 4, Appl1
11	1233.5	35.2	1359	14	US-10-059-911-30	Sequence 30, Appl1
12	1284	35.0	1422	14	US-10-059-911-26	Sequence 26, Appl1
13	1185	33.8	1231	12	US-10-425-114-11565	Sequence 11565, A
14	1171.5	33.5	1949	12	US-10-424-599-123985	Sequence 123455, A
15	1129.5	32.3	1113	14	US-10-059-911-32	Sequence 32, Appl1
16	1042	29.8	927	14	US-10-059-911-27	Sequence 27, Appl1
17	977	27.9	7367	13	US-10-024-432-7	Sequence 7, Appl1
18	974.5	27.8	1466	12	US-10-425-114-3280	Sequence 3280, Ap
19	955.5	27.3	834	14	US-10-059-911-28	Sequence 28, Appl1
20	953.5	27.2	2128	15	US-10-374-780A-1733	Sequence 1763, Ap
21	951.5	27.2	1818	15	US-10-374-780A-333	Sequence 333, Ap
22	950.5	27.1	831	14	US-10-059-911-42	Sequence 42, Appl
23	947.5	27.1	618	14	US-10-059-911-34	Sequence 34, Appl
24	945.5	27.0	1500	15	US-10-374-780A-395	Sequence 395, Ap
25	943.5	26.9	2116	12	US-10-424-599-100930	Sequence 100930, Ap
26	940	26.8	1551	12	US-10-424-599-123483	Sequence 123483, A
27	939	26.8	2088	15	US-10-374-780A-1738	Sequence 1738, Ap
28	930	26.6	2052	15	US-10-374-780A-2495	Sequence 2495, Ap
29	928	26.5	2169	12	US-10-424-599-91337	Sequence 91337, A
30	911.5	26.0	2147	12	US-10-424-599-70525	Sequence 70525, A
31	903	25.8	951	14	US-10-059-911-40	Sequence 40, Appl1
32	895	25.6	1479	12	US-10-425-114-14859	Sequence 14859, A
33	891.5	25.5	1441	15	US-10-374-780A-345	Sequence 345, Ap
34	888	25.4	1385	12	US-10-425-114-783	Sequence 783, Ap
35	887.5	25.3	1382	12	US-10-425-114-19354	Sequence 19324, A
36	869.5	24.8	794	12	US-10-425-114-90762	Sequence 20762, A
37	869.5	24.7	805	15	US-10-374-780A-1742	Sequence 1742, Ap
38	861	24.6	525	14	US-10-059-911-36	Sequence 36, Appl
39	762.5	21.8	1074	14	US-10-059-911-31	Sequence 31, Appl
40	695	19.8	1410	12	US-10-424-599-91850	Sequence 91850, A
41	695	19.8	1410	12	US-10-425-114-9718	Sequence 9718, Ap
42	695	19.8	1410	12	US-10-412-659B-1451	Sequence 9718, Ap
43	695	19.8	1410	15	US-10-374-780A-1303	Sequence 1303, Ap
44	688.5	19.7	2042	12	US-10-412-659B-1462	Sequence 1462, Ap
45	688.5	19.7	2042	15	US-10-374-780A-1309	Sequence 1309, Ap

## OTHER INFORMATION:

US-10-024-632-1

## Alignment Scores:

Pred. No.: 0  
 Score: 3502.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 13

Length: 2344  
 Matches: 663  
 Conservative: 0  
 Mismatches: 0  
 Indels: 0  
 Gaps: 0

US-10-024-632-2 (1-663) x US-10-024-632-1 (1-2344)

1 MetLysArgIleAsnGluSerAsnAsnThrAspAspGlyAsnAsnHisAsnTrpLeuGly 20  
 242 ATGAAGCGCATTAATGAGATGATACACACCGATGATGAAACATCATTAATCTGGTGGG 301  
 21 PheSerLeuSerProHisMetLysMetGluAlaThrSerAlaAlaThrValProThrThr 40  
 302 TTCTCTCTCTCAACCCACATGAAATGAGGCTACTTCAGCAGCCACTGTTCGACACAC 361  
 41 PheTrpMetSerProSerGlnSerHisLeuSerAsnPheGlyMetCysTrpGlyValGly 60  
 362 TTCTCAATGTCCTCG 421  
 61 GluAsnGlyAsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeu 80  
 422 GAAATAGGTAACTTCTCACTTCTCACTTCTCACTTCTCACTTCTCACTTCTCACTTCT 481  
 81 CysIleLeuGluAlaLeuLysArgSerGlnThrGlnValMetValProThrSerSerPro 100  
 482 TGTATCTTGGAGCTCTCAAAAGATCAACAAAGCAAGTGGTGGCCAACTTCGTCCG 541  
 101 LysLeuGluAspPheLeuGlyValAlaThrMetGlyThrHisGlyTrpGlySerHisGlu 120  
 542 AATATGGAGGACTTCTCTAGGTGTGCAACTATGGAGACTCAAGATATGAGAGCCACAG 601  
 121 ArgGlyLeuSerLeuAspSerIleTrpTrpAsnSerGlnAsnAlaGluAlaGlnProAsn 140  
 602 AAGAGTTTGAAGCTTGAACAGCATCTATATATATATATATATATATATATATATAT 661  
 141 AsnAspLeuLeuSerGlnProPheArgGlnGlnGlnHisMetSerValGlnThrHisPro 160  
 662 AAGAGCTTCTTCTTCAACACCTTCAAGGCAACAGGTCATATGATGTGCCAACACACCT 721  
 161 TrpTrpSerGlyLeuAlaCysHisGlyLeuTrpAlaProLeuGluGluGlnThrThr 180  
 722 TATTAATCTCAAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781  
 181 LysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGluGlyLeuLys 200  
 782 AAGGAAACCCACACTGTGCGATGTCACCTCCCTAATCCCTCAATATGACAGAGGCTTGA 841  
 201 AsnTrpValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlnGlnMet 220  
 842 AACTGGGTGCTTCCACAGAGGAGTTTCACTTCACTCAACAGAGTTTGGAGCAGCAATG 901  
 221 AsnCysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGlyCysGlyGlu 240  
 902 AATTTGCGCATGGGGAATGAGAAATGGTGTCTTTAGGATCTGTGGGGGTGGAGAG 961  
 241 LeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValAlaThrAlaPro 260  
 962 TTACAGCTCTAAGCTTAT 1021  
 261 SerGlyThrAspSerValAlaValAlaAspAlaLysArgGlyHisAlaLysLeuGlyGln 280  
 1022 TCTGGAAACAAATCTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1081  
 281 LysGlnProValHisArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTrpArg 300  
 1082 AAGCAGCTGTGCTATAGAAATATATATATATATATATATATATATATATATATATAT 1141

301 GlyValIleThrArgHisArgTrpThrGlyArgTrpGluAlaHisLeuTrpAspAsnSerCys 320  
 1142 GGTGTCAAGGCGATAGTGGACTGTGATGATATAGAGCCATTTGTGGATATATGTTGC 1201  
 321 LysLysGluGlyGlnThrArgLysGlyArgGlnValTrpLeuGlyGlyTrpAspMetGlu 340  
 1202 AAGAGGAAAGGCGCAACTAGAGAAAGGACGACAGTGTATTTGGGGGCTTATGATATGAG 1261  
 341 GlyLysAlaAlaArgAlaTrpAspLeuAlaAlaLeuLysTrpTrpGlyProSerThrHis 360  
 1262 GAGAAAGCTGCAAGGCTTATGATCTCGGCGCTTAAATGATCTGGGAGCTTCAACGAT 1321  
 361 IleAsnPheSerIleGluAsnTrpGlnValGlnLeuGluGluMetLysAsnMetSerArg 380  
 1322 ATTAATCTTTCATATGAGATATTCACAACTTCACTTGAAGAAATGAAATGATGACAGA 1381  
 381 GlnGlyTrpValAlaHisLeuArgArgLysSerSerGlyPheSerArgGlyValAsnTrpIle 400  
 1382 CAGGAATAGTTGCACACTTGTGAAGAAAGAAAGCAGCGGGTTCCTTGAAGGCTTCAATA 1441  
 401 TrpArgGlyValThrArgHisHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgVal 420  
 1442 TACAGAGGGGTACAAAGGATCAACATGCAATGCAAGATGCAAGCAGATGAGCAGAGT 1501  
 421 AlaGlyAsnLysAspLeuTrpLeuGlyThrPheSerThrGlnGluAlaAlaGluAla 440  
 1502 GCTGGGAAACAAAGACTTATACCTTGGAGCTTCAACACCCACAGAGAGAGCAGCAACGA 1561  
 441 TrpAspValAlaAlaIleLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSer 460  
 1562 TACGATGAGCGCGATCAAAATTTGCGCGCAAAATGCAATGCAACAACTTGAATTTCA 1621  
 461 ArgTrpAspValGluArgIleMetLysSerSerAsnLeuLeuAlaGlyGluLeuAlaArg 480  
 1622 AGATACGATGTGAGAGATCATGCGCAGTACCAATCTCTCGTGGGAGCTTCAACAG 1681  
 481 ArgLysLysAspAsnAspProArgAsnLysAspIleAspTrpAsnLysSerValThr 500  
 1682 CGTAAAGAAATACATCTTGAACAAAGGACATGACTTACAAACAGCTTGAATGAC 1741  
 501 SerValAsnAsnGluGluThrValGlnValGlnAlaGlyAsnAsnAsnGluAsnAsp 520  
 1742 AGTGTACACATGAGGAAAGGCTTCAAGTCAAGCAGGAAACACATATATGAAACGAC 1801  
 521 SerGluTrpLysMetValLeuPheAsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsn 540  
 1802 TCAGAGTGGAAAGATGTTTATTTTAAACACCTTCAACAGCAGCAACGCAATGCAAT 1861  
 541 GlySerAspGlnLysIleMetLysCysGlyAsnTrpArgAsnSerLysPheSerMetAla 560  
 1862 GGCAGTACCAAAATTAATGAACTGTGAAATTAAGAAACAGTGCATTTCTTAGGCC 1921  
 561 LeuGlnAspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlu 580  
 1922 CTACAAATCTTATTTGGATGATTCGTGGGTTCGGGAGAGATATATATATATATATAT 1981  
 581 SerSerLysIleGlyThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSer 600  
 1982 TTAGCAAAATTTGGATCTTTTCAACACGTCATCGTGTGACCAAGTTTAAACGAGC 2041  
 601 SerArgGluLysSerProGluLysArgLysProSerLeuLeuPheProMetProPheMet 620  
 2042 TCAGAGAGGCTAGTCTGAGAAAGGAGTCCCTCGCTTCTGTTCCCAAGCTTCAATG 2101  
 621 GluThrLysAlaIleValAsnProIleGlyThrSerValThrSerTrpLeuProSerProThr 640  
 2102 GAAACAAAGATGTGAACCCCATTTGATCAAGTATACCTTGGCTTACCTTCAACACAG 2161  
 641 ValGlnMetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTrp 660  
 2162 GTTCAATATGAGGCTTCTCGGATATCTTTGTGCTCACTTGCACAGTTTGTGCTTCTGG 2221  
 661 ThrAspThr 663

Db 2222 ACTGATACT 2230

RESULT 2  
US-10-424-599-136491

```

1 Sequence 136491, Application US/10424599
2 Publication No. US20040031072A1
3 GENERAL INFORMATION:
4 APPLICANT: La Rosa Thomas J
5 APPLICANT: Kovalic David K
6 APPLICANT: Zhou Yihua
7 APPLICANT: Cao Yongwei
8 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
9 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
10 FILE REFERENCE: 38-21(53223)B
11 CURRENT APPLICATION NUMBER: US/10/424,599
12 CURRENT FILING DATE: 2003-04-28
13 NUMBER OF SEQ ID NOS: 285684
14 SEQ ID NO 136491
15 LENGTH: 2405
16 TYPE: DNA
17 ORGANISM: Glycine max
18 FEATURE:
19 OTHER INFORMATION: Clone ID: PAT_MRT3847_9425C.1
20 US-10-424-599-136491

```

[illegible]

Db 1339 GAGGAGAGACACAAAGAAACGACGTT-----TCAGTAATGCGCTCAATG 1386

QY	196	Thr----GlnGluLeuYsaAntPrAlAlProCThArGluYsaPheSerThiNHisGlnGln	214
Db	1387	ACGAGAGAGAGATTGCAAAATCGGGTAGCTCTCAACAAGGAGATTCATCTACATCGACG	1446
QY	215	ValLeuGlnGlnImeIaenCySglYmeGlyAsnGluNrgAsnGlyValSerLeuGly	234
Db	1447	ATTCTGAGACGCAAAATGAACCTGGCATCTGGAAATGAGAGAACTGGGGATATCTGTGGA	1506
QY	235	SerValAlGlySglGlyLeuGlnSerLeuSerLeuSerLeuSerProGlySerGlnSer	254
Db	1507	ACTGTGGGCTGTGGAGAGTTGCATCTTAAGCTTAATCTATGAGTCTCTGTTCTCACTCT	1566
QY	255	SerCyValThralAProSerGlyThraPseValAlaValAspAlaValYbaVgLy	274
Db	1567	AGTGTGTCACTGCTCTCTTGGAAACGATTCGTGGCTGTGGATGCAAAGAGAGAGGG	1626
QY	275	HisAlaValLeuGlyGlnYsgInPProValHisArgYseSerLeaSPThrPheGlyGln	294
Db	1627	CATCTAAACTTGCTGAGAGACGCTGTGCATGAATAATCTATTGCACATTTGGGCA	1686
QY	295	ArgThSerGlnThraPrgValThraPrgHisArgTrpTrpGlyArgGlyAlaHis	314
Db	1687	AGAACGTGGAGATAGAGGCGGTCAAGGCAAGATAGAGCTGTAGTGTAGGAAGGCAT	1746
QY	315	LeuTrpaPsaPseCyValYsgInGlyGlnThraArgSglYarGlnValYrLeu	334
Db	1747	TTGGGGATATATGTGGCAAGAGAGAGGCAAGACATAGAAAGAGACAAAGATATTTG	1806
QY	335	GlyGlyThraPseGlnGluYsaAlaAlaArgAlaThraPsePheAlaIleuValSTY	354
Db	1807	CGAGCTTATATATGAGAGAAAGCTCGAAGGCTTAATGATTCGGCGCTCTTAAGTAC	1866
QY	355	TrpGlyProSerThiHisIleAsnPseSerLeuAsnTyGlnValGlnLeuGlnGlu	374
Db	1867	TGGGACCTTCACAGCAATAACTTTGCATGAAATTAACCAAGTTCAAAAAA	1926
QY	375	MetIleAsnPseSerArgGlnGlyValAlaHisLeuArgArgYseSerGlyPhe	394
Db	1927	AAAAAAMAAAAAAGCGCGCTGTCACACTTGAGAGAAAGAAAGCAGCGGTTT	1986
QY	395	SerArgGlyAlaSerIleTyArgGlyValThraHisHisGlnHisGlyArgTrpGln	414
Db	1987	TCFAGAGGTCTTCATATACAGAGGGGTCAAGGCAATCACCAATGGAATGGCAA	2046
QY	415	AlaArgIleGlyArgValAlaGlyAsnYsaPseLeuTyLeuGlyThraPseSer	432
Db	2047	GCGAGATAGGACAGATTGGTGGACAAGAACCTTTACTTGGACGTTCAAGACCCAA	2106
QY	432	-----	432
Db	2107	GAGAAAGCAGACAGACATACGATGAGCGCCGATAAATTTCCGGGCCCAATGTGTCA	2166
QY	432	-----	432
Db	2167	CCAACCTTCACACTTCAGCAACGATGTGAGAGAGCAGGCGAGTTGCATCTCTCG	2226
QY	433	-----	433
Db	2227	CTGGGGTGGCGGAGGCGCATGTGGAGATGAGAGCTAGCTCCAGGGCGCTTAGGCACC	2286
QY	434	GlnGlnGlnValAlaGlnValSTYArgValAlaAlaIlePhePheArgGlyAlaAsnAla	453
Db	2287	CAAGAGAGAGACAGACATACGATGAGCGCGCATTAATTTGCGCGGCAAAATGCA	2346
QY	454	ValThraSPheAspIleSerArgTyArgPValGluArgIleMetAlaSerSerIen	472
Db	2347	GTCACAAACTTGACATTCAGATACGATGTGAGAGAGATCATGGCCAGTAGCAAT	2403

RESULT 3  
US-10-425-114-18711

```

/ Sequence 18711, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 18711
/ LENGTH: 1378
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3109-011-P8_FLI
/ US-10-425-114-18711

```

```

Alignment Scores:
Pred. No.: 1,378-178 Length: 1378
Score: 1744.00 Matches: 331
Percent Similarity: 93.26% Conservative: 15
Best Local Similarity: 89.22% Mismatches: 13
Query Match: 49,80% Indels: 12
DB: 12 Gaps: 4

```

US-10-024-632-2 (1-663) x US-10-425-114-18711 (1-1378)

```

QY 1 MetLysArgIleAsnGluSerAsnGlnThrAspAspGlyAsnAsnHisAsnTrpLeuGly 20
DB 283 ATGAGCGCATGATGAGATGAAACACCGATGATGGAACCAATCATCATGCTGGG 342
QY 21 PheSerLeuSerProHisMetLysMetGluAlaTrpSerAlaIleThr----- 36
DB 343 TTCTCTCTCTCGGCCACATGAAATGAGGTTACTCTGGCCACCGACTGTTCCGACAC 402
QY 37 ---ValProThrThrPheThrMetSerProSerGlnSerHisLeuSerAsnPheGlyMet 55
DB 403 AAGCTCCGACCAACCTTCTCAATGTCCT-----TTCACATGTCACATCCGGAATG 456
QY 56 CysTrpGlyValGlyGluAsnGlnPheHisSerProLeuThrValMetProLeuLys 75
DB 457 TGTAAAGTGTGGGAAATGATGTAATCTTCATTCCTCTTACCGTTATGCTCTCAAG 516
QY 76 SerAspGlySerLeuGlnLeuGlnAlaLeuLysArgSerGlnThrGlnValMetVal 95
DB 517 TCGAAGGGTCATCTTGATCTTGGAGCTCTCAATGATCAACAACCAAGTATGCTG 576
QY 96 ProThrSerSerProLysLeuGlnLysPheLeuGlnGlyAlaThrMetGlyThrHisGlu 115
DB 577 CCAACTTCGCTCCGAAATGGAGGACTCTTAAGGAGGACACTATGGAACTCAGCA 636
QY 116 TyrGlySerHisGlnArgGlyLeuSerLeuAspSerIleTyrTyrAsnSerGlnAsnAla 135
DB 637 TATGGGACACGACGAGAGAGGTTGAGCTTGAACAGATCTATATATCAACAAGCA 696
QY 136 GlnAlaGlnProAsnArgAspLeuSerGlnProPheArgGlnGlnGlyHisMetSer 155
DB 697 GAGGCTCAACCCACGAAACCTTCTTCAATCCCTTAGGCAACAAGGAGTGTGAT 756
QY 156 ValGlnThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeu 175
DB 757 GTCAAAACACACCTTATCTCTGATTTGCTTGCAGGTTTATCAAGCACCGTGC 816
QY 176 GlnGlnGlnThrThrHisGlnThrHisValSerAspCysSerSerLeuMetProGlnMet 195
DB 817 GAGAGAGACCAACAAAGGAAACGCACTT-----TCAGTAAATGCTCAATAG 864

```

```

QY 196 Thr---GlnGlyLeuLysAsnTrpValAlaProThrArgGluPheSerThrHisGlnGln 214
DB 865 ACAGAGAGAGGTTTGGCAAACTGGGTAGCTTCCAAACAGAGGATATTCATCTACAGCAG 924
QY 215 ValLeuGlnGlnGlnMetAsnCysGlyMetGlyAsnGlnArgAsnGlyValSerLeuGly 234
DB 925 ATCTGGACGACCAATGAACTGTGGCATTTGGATGAGAGAGAGTGGGTATCTGTTGCA 984
QY 235 SerValGlyCysGlyGlnLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSer 254
DB 985 ACTGGGGGCTGGAGAGTTCCAACTCTTAAGCTTATCTATGAGTCTCGGCTTCACTCT 1044
QY 255 SerCysValThrAlaProSerGlyThrAspSerValAlaValAspAlaLysValArgGly 274
DB 1045 AGTTGTGACCTGCTCTCTTCTGGAAACAGATTTCTGTGCTGTGATGCAAGAGAGAGG 1104
QY 275 HisAlaLysLeuGlnGlnLysGlnProValHisArgLysSerIleAspThrPheGlyGln 294
DB 1105 CATGCTAAACTTGTGTGACAGACACCTGTGATGAGAAATCTATTTGACACATTTGGGCAA 1164
QY 295 ArgThrSerGlnTyrArgGlyValThrArgHisArgTrpThrArgValTyrGlnAlaHis 314
DB 1165 AGACGCTGCAGATATGAGGCGTCCACAGGCAATGATGATGATGATGATGATGATGAT 1224
QY 315 LeuTrpAspAsnSerCysLysLysGlnGlyGlnThrArgLysGlyArgGlnValTyrLeu 334
DB 1225 TTGTGGATTAATATGATTGCAAGAGAGAGGACGACCTAGAAAGAAAGAAATATTTG 1284
QY 335 GlyGlyTyrAspMetGlnGlnLysAlaAlaArgAlaTyrAspLeuAlaLeuLysTyr 354
DB 1285 GGAGTTTATGATATGAGAGAAAGAAAGCTGCAAGAGACCTTATGATCTCGGCTTTAAGTAC 1344
QY 355 TrpGlyProSerThrHisIleAsnPheSerIle 365
DB 1345 TGGGAGCTTCAACGACATTAACCTTTGATA 1377

```

## RESULT 4

US-10-024-632-3

```

/ Sequence 3, Application US/10024632
/ Publication No. US20020170093A1
/ GENERAL INFORMATION:
/ APPLICANT: Monsanto Technology LLC
/ APPLICANT: He, Steve S.
/ APPLICANT: Dolson, Stanton B.
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION
/ TITLE OF INVENTION: GROWTH AND USES THEREOF
/ FILE REFERENCE: 38-21(51857)B
/ CURRENT APPLICATION NUMBER: US/10/024,632
/ PRIOR APPLICATION NUMBER: US 60/257,896
/ PRIOR FILING DATE: 2000-12-21
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 323
/ LENGTH: 2323
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (25)..(2022)
/ OTHER INFORMATION:
/ US-10-024-632-3

```

```

Alignment Scores:
Pred. No.: 4,246-175 Length: 2323
Score: 1715.00 Matches: 378
Percent Similarity: 65.86% Conservative: 83
Best Local Similarity: 54.00% Mismatches: 165
Query Match: 48.97% Indels: 74
DB: 13 Gaps: 20

```

US-10-024-632-2 (1-663) x US-10-024-632-3 (1-2323)

QY 1 MetIyArGllEaSnGlSeRnSbNthRAsPArGlyAaSnRnHsAeNtPLeuGly 20  
 DB 25 ATGAAGAGATGTGAAATGATGACAATGCTGCTTAATATCAAAACAAATGTTGGGT 84  
 QY 21 PheSerLeuSerProHsEmet-----LysMetGluAlaThr 32  
 DB 85 TTCTCACTCTCTCCCAATGCAATAATAGAGTTTCTGCACTCAACACTTCCCT 144  
 QY 33 SerAlaAlaThrValProThrThrPheTyMetSerProSerGlnSerHsLeuSerSn 52  
 DB 145 GCTGCTGAAGTGGTCTCTCAAGCTTTTACCAACACACTGCTCA-----CTTAGTAC 198  
 QY 53 PheGlyMetCysTyGlyVal---GlyGluAaNGlyAenPheHsSerProLeuThrVal 71  
 DB 199 TATGTTTCTACTATGAGCTTGAAGCTGAAAATGTTGATTTGTATTCAGCTTTGCCAATC 258  
 QY 72 MetProLeuLysSerApsGlySerLeuCysIleLeuGluAlaLeuLysArgSerGlnThr 91  
 DB 259 ATGCCCCCTCAAACTGATGGCTCTCTCTATGATTGGAACCTTTAGCGAGTCAAGCA 318  
 QY 92 GlnValMetValProThrSerSerProLysLeuGluAaSPheLeuGlyGlyAlaThrMet 111  
 DB 319 CAAGCAATGGCTACTCTTCAACACCAAACTGGAGAACTTTTAGTGGGGAAGCCATG 378  
 QY 112 GlyThr---HisGluTyT-----GlySerHisGluArgGlyLeuSerLeuApsSer 127  
 DB 379 GGGACCCCTCATCATGACCAATGTAGTGCACAGAAACAATGCTCTGAGCTTGAACAT 438  
 QY 128 IleTyTyTrAsnSerGlnAaSnIaGlyAlaGlnProAaSnArgApsLeuSerGlnPro 147  
 DB 439 GTTTTATAC---ATCCAACTCAGCCGCTGACCCCAATATATATCAAACTTACCAAAAC 495  
 QY 148 PheArgGln-----GlnGlyHisMetSerValGlnThrHisProTyTyTr 162  
 DB 496 CATGTTCAACACATAGACCAACCAACCAACCAACCAAGAGCTTCAAGCATATATAC 555  
 QY 163 SerGlyLeuAlaCysHisGlyLeuTyGlnAlaProLeuGlnGluGlnThrHisGly 182  
 DB 556 TCTACCTTGAAGAACCATGATATGATATTA-----GAAAGGTCAAAGCA 600  
 QY 183 ThrHisValSerApsCysSerSerLeuMetProGln----- 194  
 DB 601 AGCCAACTTCTGCAACCAACATCTTCACTTCAAAACATGGGTGTATATGCGCGTT 660  
 QY 195 MetThrGluGlyLeuLysAsnTPValAlaProThrArgGluPheSer-----Thr 211  
 DB 661 CCGTTCCTGGCCCTCAAGAGTTGG-----GAAGTGAGAACTTCCAAAGCTTACGCA 714  
 QY 212 HisGlnGln-----ValLeuGlnGlnGlnMetAaSnCysGlyMetGlyAaNGlyAa 228  
 DB 715 CATAGACCAAGATGATTTGCTCATGTGAGAAAATGCTGTGAATAGAG----- 768  
 QY 229 AsnGlyValSerLeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeuSerMet 248  
 DB 769 -----TCCATTGATCAATGAGCTTATGCTGACCTTGCATGTTGAGCTTGCATG 819  
 QY 249 SerProGlySerGlnSerSerCysValThr-----AlaProSerGlyThr 263  
 DB 820 AGTCTACTCTCTCACTTGAAGAGTGTCAAGATTCTCACCGTGGCTTCACTGCTGTGTT 879  
 QY 264 ApsSerValAlaValAspAlaLysLeuArgGlyHisAlaLysLeuGlyGlnLysGlnPro 283  
 DB 880 GATTCTGTTGCCATGATATCTAAGAAAAGGGGCTGAAAAGCTTGAACAGAACAAAT 939  
 QY 284 ValHisArgLysSerIleAspThrPheGlyGlnAlaGlnThrSerGlnTyArgGlyValThr 303  
 DB 940 GTTCAATAGGAAGTCCATGACACCTTTGCAAAAGAACCTTCCAGTATAGAGAGCTAAC 999  
 QY 304 ArgHisArgGlyTrpThrGlyArgTyGlyAlaHisLeuThrApsAaSnSerCysLysLysGln 323  
 DB 1000 AAGGCAATAGGTGACTGGAGATATATAGCTCATCTTTGGGCAACAGCTGCAAGAAAGAG 1069  
 QY 324 GlyGlnThrArgLysGlyValArgGlnValTyLeuGlyGlyTyTrApsMetGluGluValAla 343

DB 1060 GGCGAAAGCGAGAAAGAAAGACAAAGTTTATCTAAGGGGCTTATGATATGGAAGAAAGCT 1119  
 QY 344 AlaArgAlaTyTrApsProLeuAlaAlaLeuLysTyTyTrGlyProSerThrHisIleAenPhe 363  
 DB 1120 GCGAGAGCTTATGATCTTACGGGACCTCAAGTATTTGGGAGCCCTCCACTACATTAACCTTT 1179  
 QY 364 SerIleGluAaNGlyGlnValGlnLeuGlnGlyMetLysAaSnMetSerArgGlnTyTr 383  
 DB 1180 CTTTGAAATATATCAAAATGCAATCTTGAAGAAATAGAAACATGACTAGACAAAGATAT 1239  
 QY 384 ValAlaHisLeuArgThrArgLysSerSerGlyPheSerArgGlyValaSerIleTyArgGly 403  
 DB 1240 GTTGCTCATTTGGAAGAAAACAGCGGATTTCTCAAGAGGGCTTCCATGTCAGAGCA 1299  
 QY 404 ValThrArgHisHisGlnHisGlyArgTyGlnAlaArgIleGlyArgValAlaGlyAa 423  
 DB 1300 GTAACAAAGACCAACCAACATGGAAGGTGGCAAGCTGGAATTGTTAGAGTGGCGAAAC 1359  
 QY 424 LysApsLeuTyLeuGlyTyTrPheSerThrGlnGlnAlaAlaGlyAlaTyTrApsVal 443  
 DB 1360 AAAGATCTATATCTTGAACCTTTAGTACCAAGAGAACAGCTGAAGCTTATAT 1419  
 QY 444 AlaAlaIleLysPheArgGlyAlaAaSnAlaValThrAaSnPheApsIleSerArgTyTrAps 463  
 DB 1420 GCTGCTATTAATTTCCAGAGAGCGAATGCTGTAAACAACTTTGACATCACAAGATATGAT 1479  
 QY 464 ValGluArgIleMetAlaSerSerAaSnLeuAlaGlyGluLeuAlaArgAlaGlyLys 483  
 DB 1480 GTGAGAAAATCATGAGCAAGAGCAACCTCTTAGCGAGTGAAGCTTACGCAAGCA 1539  
 QY 484 ApsAaSnProCysAaSnLysApsIleApsTyTrAaSnLysSerValValThrSerVal--- 502  
 DB 1540 GAGAGCGCAATGAACCTCACTGATGATTAATATGACAAATAGCTTGTGCATATGAG 1599  
 QY 503 AsnAaNGlnGluThrValGlnValGlnAlaGlyAaSnAaSnAaNGluAaSnApsSerGlu 522  
 DB 1600 GACACTGAAGAGCATCTTAATGACACAGAAAGCTGTGAGAGCAAAATGAT---CAG 1656  
 QY 523 TrpLysMetValIlePheAaSnHisProSerGlnGlnGlnAlaAaNGlyAaNGlySer 542  
 DB 1657 TGGAGAGTGTCTCTAC---CAATCCTGAGCACTTGAAGCAAAATCCCAACAAT 1713  
 QY 543 ApsGlnLysIleMetAaSnCysGlyAsnTyTrArgAaSnAlaPheSerMetAlaLeuGln 562  
 DB 1714 CAG-----AGTGCAGAACTAACCAAGCTCTTGCAGTGGCTTTGGAC 1755  
 QY 563 ApsLeuIleGlyIleApsSerValGlySerGlyGlnHisAaSnMetLeuApsGluSerSer 582  
 DB 1756 AACCATGTT-----CATCAGAAATGAGAGATCATAGT 1788  
 QY 583 LysIleGlyThrHisPheSerAaSnThzSerSerLeuValThrSerLeuSerSerArg 602  
 DB 1788 AAGCGAGAGCCATGTCTCAATTCCTTCACTTATGGGCCACAAGTTTGAAGCGCTCAAGA 1848  
 QY 603 GlnAlaSerProGlnLysArgGlyProSerLeuLeuPheProMetProPrometGlnThr 622  
 DB 1849 GAAGGTAGCCCTGATAGCAAGACTTGGCAATGCTCTGGAATCGCTTCAACTGCATCA 1908  
 QY 623 LysIleValAaSnProIleGlyThzSerValThzSerTrpLeuProSerProThrValGln 642  
 DB 1909 AAACATTTGGCTACTAATCCAAATACATGAATTTTGGAGCCCTTCAACCCCATTTGAGG 1968  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisIleProValPheAlaSerTrpThrSer 662  
 DB 1969 -----CCAGCACTTACTTGTGCTTCAATGCAAGTTTGGAGCTTGAAGAT 2016

RESULT 5  
 US-10-024-632-5  
 ; Sequence 5, Application US/10024632  
 ; Publication No. US20020170093A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Monsanto Technology LLC

APPLICANT: He, Steve S.  
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION  
 TITLE OF INVENTION: GROWTH AND USES THEREOF  
 FILE REFERENCE: 38-21(51837)B  
 CURRENT APPLICATION NUMBER: US/10/024,632  
 PRIOR FILING DATE: 2001-12-19  
 PRIOR APPLICATION NUMBER: US 60/257,896  
 NUMBER OF SEQ ID NOS: 33  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO: 5  
 LENGTH: 1926  
 TYPE: DNA  
 ORGANISM: *Oryza sativa*  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)..(1926)  
 OTHER INFORMATION:  
 US-10-024-632-5

Alignment Scores:  
 Pred. No.: 5,01e-142 Length: 1926  
 Score: 1408.50 Matches: 340  
 Percent Similarity: 55.05% Conservative: 63  
 Best Local Similarity: 46.45% Mismatches: 152  
 Query Match: 40.22% Indels: 177  
 Gaps: 26

US-10-024-632-2 (1-663) x US-10-024-632-5 (1-1926)

QY 13 GYAsnAenHiasnTPrLeuGlyPheSerLeuSerProHismetLysMetGluAlaThr 32  
 DB 13 GGGCGACAGAGAACTGGTTAGGCTTCGCTCCCGACATGCCGCCATGSAAGTG 72  
 QY 33 SerAlaAlaThrValProThrThr----- 40  
 DB 73 CCGTCTCTCTGACCCATGACATGCTGCTCATCATCATCATCATCATCATCATCTGCT 132  
 QY 41 -----Phe 41  
 DB 133 GGTGCTGCTGCTCCGCGACCATGTCCTCTCTCCGACAGCGCCAGACCTGCACTTC 192  
 QY 42 TyrMetSerProSerGlnSerHisLeu-----SerAsnPhgGly 54  
 DB 193 CTCTTCTCTCTCTGACGACAGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 252  
 QY 55 MetCysTyrGlyValGlyGluAsn-----GlyAsnPhenHisSerProLeuThrValMet 72  
 DB 255 GGGCGCTACGAGAGGAGGACGACGACCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 312  
 QY 73 ProLeuLysSerAspGlySerLeuCysIleLeuGluAlaLeuLysArgSerGlnThrGln 92  
 DB 313 CTTATCAAGTCCGATGGCTCCCTCTGACATCATGGAAGC----- 351  
 QY 93 ValMetValProThrSerSerProLysLeuGluAsnPhenGlyValAlaThrMetGly 112  
 DB 352 ---ArgATGCCCTGCTCATGCGCAAGCTCGAGACTCTTGGGGTGGCGAAAGGCACT 408  
 QY 113 ThrHisGluTyrGlySerHisGluArgGlyLeuSerLeuAspSerIleTyrTyrAsnSer 132  
 DB 409 GGGCAT-----GACCGCGCCACCTACTATATAGC 435  
 QY 133 GlnAsnAlaGluAlaGlnProAsnArgAspLeuLeuSerGlnProPhaArgGlnGly 152  
 DB 436 CAGGGCCAAAGAGCGAGGAT-----GCAAGCAGGGCGGCTTACAGAC 486  
 QY 153 HisMetSerValGlnThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGln 172  
 DB 481 CACCAAGCTAGTC-----CCCTACAAC-----TACAG 507  
 QY 173 AlaProLeuGluGluGluThrThrHisGluThrHisValSerAspCysSerSerLeuMet 192

DB 508 CCAATGACGAAGAGAGATGCTGCAAGAGCGCCAGCGGCCAATGAGAGCAATG 567  
 QY 123 ProGlnMetThrGluGluLeuLysAsnTrpValAlaProThrArgGluPheSerThrHis 212  
 DB 568 -----CGGGCGGCGCAAGAACTTC----- 585  
 QY 213 GlnGlnValLeuGluGlnGlnMetAsnGlyMetGlyAsnGluArgAsnGlyValSer 232  
 DB 586 -----CTGTCACAGCTACGCGCGCTGCTACGCGCAACAGAGATGCGC----- 630  
 QY 223 LeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySer 252  
 DB 631 -----CAGCGCTCAGCCTCTCTCATATAGCGCGCTCC 663  
 QY 253 Gln---SerSerCysValThr---AlaProSerGlyThrAspSerValAlaValAspAla 270  
 DB 664 CAGTCCAGCAGCTGCTGATGCACTGCCAGACGATCAGACAGATGCGGTGGTGGCT 723  
 QY 270 ----- 270  
 DB 724 GCAGCTGCTGCTGCTGATGCGCGAAGCAACAGTAATGACGTGGCGAGCAGCT 783  
 QY 271 -----LysLysArgGlyHisAlaLysLeuGlyGlnLysGlnProValHisArgLysSer 288  
 DB 784 GTGGGAAAGAGAGGCGACCGGAAAGGCGCCAAAGACAGCTTTCACCGAAGTCC 843  
 QY 289 IleAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValAlaThrArgHisArgTrpThr 308  
 DB 844 ATTGACAGCTTGGGACAGAGACATCGCATATAGGGCGCTCACCGACAGCAGGTGGACT 903  
 QY 309 GlyArgTyrGluAlaHisLeuTrpAsnAsnSerCysLysLysGluGlyGlnThrArgLys 328  
 DB 904 GGAAGATATGAAGCCCACTCGGATACAGTTCGCAAAAGAGATGACACAGCAAGAG 963  
 QY 329 GlyArgGlnValTyrLeuGlyGlyTyrAspMetGluGluValAlaAlaArgAlaTyrAsp 348  
 DB 964 GGAAGCAAGTATCTAGTGGTTATGACACTGAAGATAAGCTGCAGAGGCTTATGAT 1023  
 QY 349 LeuAlaAlaLeuLysTyrTrpGlyProSerThrHisIleAsnPhenSerIleGluAsnTyr 368  
 DB 1024 CTGGCTGCGCTGAATACTGGGGGCTATCTACGCTAATAATTTCCGTTGAATACTAC 1083  
 QY 369 GlnValGlnLeuGlnGlnMetLysAsnMetSerArgGlnGluTyrValAlaHisLeuArg 388  
 DB 1084 CAGATGAGATGAGAGATGGAAGATGACAAAGCAAGAAATATCTTGGCACTTGA 1143  
 QY 389 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValAlaThrArgHis 408  
 DB 1144 AGGAGAGAGAGGGGTTCTCTCGCGGTCTTCATCTACCGGGAGTAAAGAGGATCAC 1203  
 QY 409 GlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeu 428  
 DB 1204 CAGATGAGATGAGAGATGGAAGATGACAAAGCAAGAAATATCTTGGCACTTGA 1263  
 QY 429 GlyThrPheSerThrGlnGlnGluAlaAlaGluAlaTyrAspValAlaAlaIleLysPhe 448  
 DB 1264 GGCATTTCAGCATCTCAAGAAAGAGAGAGGAGATGACATTGCTGCCCTCAAGTTC 1323  
 QY 449 ArgGlyAlaAsnAlaValThrAsnPhaPheAspIleSerArgTyrAspValGluArgIleMet 468  
 DB 1324 CGTGGCTGAACGGGTGCGAAGCTTTGACACACAGAGTACAGCTGACAGACATCATG 1383  
 QY 469 AlaSerSerAsnLeuLeuAlaGlyLeuLeuAlaArgArgLysValys----- 483  
 DB 1384 GAGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1443  
 QY 484 ---AspAsnAspProArgAsnLysAspIleAspTyrAsnLysSerValValThrSerVal 502  
 DB 1444 CCGGACCATGCTCAATAGCGCGAGCTC-----GCTGGG 1479  
 QY 503 AsnAsnGluGluThrValGlnValGlnAlaGlyAsnAsnAsnAsnGluAsnAspSerGlu 522  
 DB 1480 ACCGAGAAAGCGAGCGCTGCTACTGTCACGGGC-----ACCGAC 1518







QY 345 rgaAtaTyrApleuAalalaLeuYrTyrGlyProSerThrHisIleAsnPheseri 365  
 Db 1055 GGGCGATGATGATCTTGGCGCTCAAGTACTGGGGGCGCTTCCACGACATCAACTCCCGT 1114  
 QY 365 legIuBantYrGlnValGlnLeuGluGluMetLeuAsnMetSerArgGlnGluTyrVala 365  
 Db 1115 TGAAGACTACCGAG 1174  
 QY 385 laHisLeuArgArgSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValT 405  
 Db 1175 CTCACCTCGAG 1234  
 QY 405 hrarGhISHISGlnHISGlyArgTyrGlnAlaArgIleGlyArgValaIleGlyLeuYsa 425  
 Db 1235 CCAAGCATCATCGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1294  
 QY 425 spleuTyrLeuGlyThrPhe----- 431  
 Db 1295 ACCCTTACTTGGGAGCATTCATCGCGCTTTCGCGCGCGCGCGCGCGCGCGCATG 1354  
 QY 432 --SerThrGlnGlnGlnAlaAlaGlnAlaTyrAspValAlaAlaIleTyrPheArgGly 451  
 Db 1355 CCGGACACGAG 1414  
 QY 451 laAsnAlaValThrAsnPheAspIleSerArgTyrAspValGluArgIleMetAlaSer 471  
 Db 1415 TCACACCGCTCGACCACTTCGACATCACAGAGTACAGAGTGCAGACAGAGAGAGAG 1474  
 QY 471 exaAsnLeuLeuAlaGlyGlnLeuAlaArgArgIle---LysAspAsnAspProArgAsn 490  
 Db 1475 GCACGCTCTCCCGGGGAG 1533  
 QY 490 ysaAspIleAspTyrAsnLysSerValValThrSerValAsnAsnGlnIleThrValGln 510  
 Db 1534 -----GCGCGCGCGCGCTCGCGAGCGCGCGCGCGCGCGCGCTTGG 1567  
 QY 510 alGlnAlaGlyAsnAsnAsnAsnGlnAsnAspSerGluTyrPheMetValLeu----- 527  
 Db 1568 TCGAGGCGCGG-----MACGTGGCGGAGTGCAGAGATGAGCCACCGCGCGCG 1612  
 QY 528 -----PheAsnHisProSerGlnGlnGlnGlnAlaAlaGlnIleAsnGlySerAspGln 545  
 Db 1613 CGCTGCAGCGCGCGAG 1671  
 QY 545 yarIleMetLeuAsnGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeu 565  
 Db 1672 -----CACCATGACCTTCCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1720  
 QY 565 legIy---IleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlnSerSerIys 584  
 Db 1721 TGTTCGACCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1755  
 QY 584 legIyThrHisPheSerAsnThrSerSerIleValThrSerLeuSerSerArgGlna 604  
 Db 1756 --CGCGCGGACATGTCGATGCGCGG-----ACAGAGCTTGCGCACTCCGGGAGC 1804  
 QY 604 laSerProGluValArgGlyPro-----SerLeuLeuP 615  
 Db 1805 AGAGCGCTGAC--AGGGGCGTCCGCGCGCGCGCGCGCGCGCGCTCTCCACAGCGGT 1861  
 QY 615 heProMetProPheMetGlnThrIleValAsnProIleGlyThrSerValThrSerT 635  
 Db 1862 TCGCGAAGCCCGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1915  
 QY 635 rpleu---ProSerProThrValGlnMetArgProSerProAla---IleSerLeuSerH 653  
 Db 1916 GGGCGCTGCGCTCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1975  
 QY 653 laLeuProValPheAlaSerTyrThrAsp 662  
 Db 1976 ACCGCGCAATGTTCCGCGCGTGCAGCCGAC 2004

US-10-059-911-1  
 / Sequence 1, Application US/10059911  
 / Publication No. US20030159180A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Fischer, Robert L.  
 / APPLICANT: Mizukami, Yukiko  
 / APPLICANT: The Regents of the University of California  
 / TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
 / FILE REFERENCE: 023070-090730US  
 / CURRENT APPLICATION NUMBER: US/10/059, 911  
 / CURRENT FILING DATE: 2002-06-17  
 / NUMBER OF SEQ ID NOS: 42  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 1  
 / LENGTH: 2148  
 / TYPE: DNA  
 / ORGANISM: Arabidopsis thaliana  
 / FEATURE:  
 / OTHER INFORMATION: AINTEGUMENTA (ANT) CDNA  
 / FEATURE:  
 / NAME/KEY: CDS  
 / LOCATION: (269)..(1936)  
 / OTHER INFORMATION: Arabidopsis AINTEGUMENTA (ANT)  
 US-10-059-911-1  
 Alignment Scores:  
 Pred. No.: 3,116-133 Length: 2148  
 Score: 1328.00 Matches: 312  
 Percent Similarity: 52.84% Conservative: 69  
 Best Local Similarity: 43.27% Mismatches: 116  
 Query Match: 37.92% Indels: 224  
 DB: 14 Gaps: 21  
 US-10-024-632-2 (1-663) x US-10-059-911-1 (1-2148)  
 QY 1 MetLeuArgIleAsnGluSerAsnAsnThrAspAspGlyAsnAsnHisAsnTyrLeuGly 20  
 Db 269 ATGAAGTCTTTTGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 328  
 QY 21 PheSerLeuSerProHis--MetLysMet----- 29  
 Db 329 TTTCTATTGCTCTCAAAATGATGATAAATGGAGAGTAGAGAGTAGAGAGAGAGAGAGAG 368  
 QY 30 -----GluAlaThrSerAlaAlaThrValProThrThrPheTyrMet 43  
 Db 389 TCATCTCACTTCATTCAGTCAAGTCACTTCCTTCCTGTTCCACTCAACTGTTGTT 448  
 QY 44 SerProSerGlnSerHisLeuSerAsnPheGlyMetCystIleGlyValGlyGlnAsnGly 63  
 Db 449 GGT-----GACACACTAGCAACTTGGTGTGCTATGATGATTAACCAATGGA 499  
 QY 64 AsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeuCyIleLeu 83  
 Db 500 GGAATCTATTCATCATGATCTGTGATGCACTGATGATGATGATGATGATGATGATGATG 559  
 QY 84 GluAlaLeuYsaProSerGlnThrGlnValMetValProThrSerSerProLysLeuGln 103  
 Db 560 GAAGCTCTCAACAGATCTTCTCACTCGAATCAACATGAATGATATCTCAAAAGGTGAG 619  
 QY 104 AspPheLeuGlyGlyAlaThrMetGlyThrHisGlnTyrGlySerHisGlnArgGly--- 122  
 Db 620 GATTTCTTGGG-----ACCATCAACAACAACAAGTCAACAAGAAGCATG 667  
 QY 123 ---LeuSerLeuAspSerIleTyrTyrAsnSerIleAsnAlaGlnAlaGlnProAsnArg 141  
 Db 668 GATCTTAGCTTAGAGATGATTAATCTTACACACCACTCAT-----GAGCCCAACG 718  
 QY 142 AspLeuLeuSerGlnProPhe-----ArgGlnGlnGlnHisMetSerValGln 157  
 Db 719 ACTACAACTTTCAAGAGATCTTAGCTCCCTCAACCAAGAAACCAT----- 766  
 QY 158 ThrHisProTyrTyrSerGlyLeuAlaCysHisIleLeuTyrGlnAlaProLeuGlnGln 177







Db	1429	GAA	-----	1431
Qy	526	ValLeuPheAenHisProSerGlnGlnGlnGlnAlaSerGlyAsnGlySerAspGlnLys		545
Db	1432	----	-----GACCCAACC	1440
Qy	546	IleMetAsnGlyGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle		565
Db	1441	GCTCTAATGCT	-----	1452
Qy	566	GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlnSerSerLysIleGly		585
Db	1452	----	-----	1452
Qy	586	ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerArgGlnAlaSer		605
Db	1453	-----GTTGCGAAGGTGCTTCCACAAAGAACTCAGT		1485
Qy	606	ProGlnLysArgGlyProSerLeu-----LeuPheProMetProProMetGlnThrLys		623
Db	1486	ACTCCGAGAGACTCTTGAGTTTCCGCGCATTTTCGCGTCCCTCAAGATATCAAAAG		1548
Qy	624	IleVal-----AsnProIleGlyThrSerValThrSerThrPleProSerProThrValGln		642
Db	1546	ATGTTCCGAGTCAAAATATGGCGCGAAATATGACTCCTTGACACTCAAAACCTTAATGCTAG		1605
Qy	643	MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTyrThrAsp		662
Db	1606	CTTAAG-----ACCGTGCCTTCACTTGCTCCTCAGATGCGCGATTTCGTGCTGGCGTGAT		1662
Qy	663	Thr 663		
Db	1663	TCT 1665		





```

OY 198 G1yleuysaentpValAlaProthrArgIuPheserThrhsgIngluValleuGlu 217
DB 544 GGA-----546
OY 218 GIngluMetanCysGlyMetGlyAsnGluArgasnIlyValserIeuGlySerValGly 237
DB 547 GGGCTTTTAAATGAGGGGTATATGGGAATTCAA-----582
OY 238 CysGlyGluLeuInserLeuSerLeuSerMetSerProGlySerGlnSerSerCysVal 257
DB 583 -----CAGTCACTGAGCTTATCCATGAGCCCTGGGTCAACAATCTAGCTGCATC 630
OY 258 ThrAlaProser-----261
DB 631 ACTGGCTCTCACACCACCAACAAACCAAAACCAACCAACCAACCAACCAAC 630
OY 262 -----GlyThrAspSerValAlaVal 268
DB 691 CAGATCTCTGAAGCTCTTGAGAGACAAGCGTTGGGTTTGAGACGACGACATGGCGGCT 750
OY 269 AspAlaIlyAspArgIyHis-----AlaIlySerGlyGlnIlyGlnProValHis 285
DB 751 GCGAAGAAAGAGGGGGAACAAGAGATGTTGATGTTGATGTTGATGTTGATGTTGAT 810
OY 286 ArgIlySerIleAspThrPhagIyGlnArgThrSerGlnIlyArgGlyValThrArgHis 305
DB 811 AGAAATCTATCGATCTTTTGAGACACGAACTTCAATCCAGGCGGTACAGACAT 870
OY 306 ArgIyPThrIyArgIyGlyGlnAlaHisLeuIlyPAspAsnSerCysIlyIyGlyGln 325
DB 871 AGATGAGCTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 930
OY 326 ThrArgIyGlyArgIyValIlyIyGlnIlyIyPAspMetGluIlyValAlaArg 345
DB 931 AGTAAAGAAAGAGACAGATTTATCTGGAGGTTTATGATGATGATGATGATGATGATGAT 990
OY 346 AlaIyAspLeuAlaIleuIlyIyIyPThrIyIyPThrIyIyIleAsnPheserIle 365
DB 991 GATATGATCTTGCTGCACTCAAGTACGAGGTCCCTCTACTCACCAATTTCTGCG 1050
OY 366 GluAsnIyGlnValGlnleuGlnIlyMetIyAsnMetSerArgGlnIlyValAla 385
DB 1051 GGAATTTATCAGAAAGAGATTTATGATGATGATGATGATGATGATGATGATGATGAT 1110
OY 386 HisLeuArgArgIySerSerIyPAspArgIyAlaSerIleIyArgGlyValThr 405
DB 1111 CATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1170
OY 406 ArgHisHisGlnHisGlyArgIyPThrIyIyIyGlyArgValAlaGlyAsnIlyAsp 425
DB 1171 AGCATCACACACAGATGAGAGGTCAGACAGCATGATGATGATGATGATGATGATGATGAT 1230
OY 426 LeuIyIyLeuGlyIyThrPheserThrhsgIngluIlyAlaGluAlaIyAspValAla 445
DB 1231 CTTCTACTTGGAACTTTTGGAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1290
OY 446 HisLeuPAspArgIyAlaAsnAlaValIlyIyIyIyIyIyIyIyIyIyIyIyIyIy 465
DB 1291 ATTAAAGTTCCTGGCACAATGCTGTACTTAATGATATCAGAGAGAGAGAGAGAGAG 1350
OY 466 ArgIle 467
DB 1351 CGTATC 1356

```

```

RESULT 12
US-10-059-911-26
; Sequence 26, Application US/10059911
; Publication No. US20030159180A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Mizukami, Yukiko
; APPLICANT: The Regents of the University of California

TITLE OF INVENTION: Methods for Altering Organ Mass in Plants
FILE REFERENCE: 023070-090730US
CURRENT APPLICATION NUMBER: US/10/059,911
CURRENT FILING DATE: 2002-06-17
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 26
LENGTH: 1422
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:truncated or
OTHER INFORMATION: mutated ANT cDNA transgene ANTDM1
US-10-059-911-26

Alignment Scores:
Pred. No.: 3,03e-122 Length: 1422
Score: 1224.00 Matches: 284
Percent Similarity: 53.46% Conservative: 56
Best Local Similarity: 44.65% Mismatches: 92
Query Match: 34.95% Indels: 204
DB: 14 Gaps: 18

US-10-024-632-2 (1-663) x US-10-059-911-26 (1-1422)
OY 69 LeuThrValMetProLeuIlySerAspGlySerLeuCysIleLeuGluAlaLeuIlyArg 86
DB 1 AGCTGCTGATCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
OY 89 SerGlnThrGlnValMetValProThrSerSerProIyIyLeuGlyAspPheLeuGlyIy 108
DB 61 TCTTCTCACTCAATCAACCATCAAGATTCATCTCCAAAGCTGAGAGATTTCTTGG 117
OY 109 AlaThrMetGlyThrHisGlyIyIyGlySerHisGlyIyGlyIyIyIyIyIyIyIyIy 126
DB 118 -----ACCATCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 168
OY 127 SerIleIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 146
DB 169 AGTTATTTTACCAACCACTCAT-----GAGCCCAACACACATCAAACTTTCA 219
OY 147 ProPhe-----ArgGlnGlnIyHisMetSerValGlnThrHisProIyIyIy 162
DB 220 GAGTTCTTAGCTCCCTCAACCAACAACAACA-----252
OY 163 SerGlyLeuAlaCysHisIleIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 182
DB 253 -----GAGGAGGAAACT-----AGA 267
OY 183 ThrIleValSerAspCysSerSerLeuMetProGlnMetThrhsgIyLeuIyAsnIyP 202
DB 268 AATTACGGGAGATAC-----CTAGTTGACACATGGA-----300
OY 203 ValAlaProThrArgIuPheserThrhsgIngluValleuGlnleuMetanCys 222
DB 301 -----GGGCTTTTAAATGTA 315
OY 223 GlyMetGlyAsnGluArgasnIlyValserLeuGlySerValGlyCysGlyGluIyGln 242
DB 316 GGGGTATATGGGAATTTCAG-----CAG 339
OY 243 SerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValThrAlaProser 261
DB 340 TCACTGAGCTTATCCATGAGCCCTGGGTCAACAATCTAGCTGCATCTGCTTCAACAC 399
OY 261 -----261
DB 400 CACCAACAAACCAAAACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 459
OY 262 -----GlyThrAspSerValAlaValAlaAspAlaIlyIyIyIyIyIyIyIyIyIy 273
DB 460 CTTGTGAGACAAGCGTTGGGTTTGAAGACGACGACGACATGGCGGTCTGGAAGAGAGAG 519

```



QY 274 GAlHs-----AlaLysLeuGlnGlnGlnProValHisArgLysSerIleAsp 290  
 DB 520 GGACAGAGAGATGTTAGTGTGTGTGAGAAAGAGATGTTCAATAGAAAATCATATGAT 579  
 QY 291 ThrPheGlyGlnArgThrSerGlnTyraGlyValIleThrArgHisArgIlePheGlyArg 310  
 DB 580 ACTTTGGACAGCAAGCTTCTCAATACAGAGCGCTTCAACAGACATAGATGACATGTTGAG 639  
 QY 311 TTTGUAAlHisLeuTrpAspAsnSerCysLysLysGlnGlnTrpArgLysGlyArg 330  
 DB 640 TATGAGGCTCATCTATAGGACATATGTTCAAGAGAGAGATGACAGTAAAGAAAAGAGA 699  
 QY 331 GlnValTyraLeuGlnGlyTyraPheMetGlnGlnValAlaAlaGlyAlaTyraPheLeuAla 350  
 DB 700 CAAGTTATCTGGAGAGTTATGATATGAGGAGAAAGCTGCTCAGACATATGATCTTGGCT 759  
 QY 351 AlaLeuLysTyraTrpGlyProSerThrHisIleAsnPheSerIleGluAsnTyraGlnVal 370  
 DB 760 GACCTCAAGTACTGGGGGCTCCCTCTACTACACCAATTTCTCTGGGAGAAATTATACAGAA 819  
 QY 371 GlnLeuGlnGlnMetLysAsnMetSerArgGlnGlnTyraValAlaHisLeuArgArgLys 390  
 DB 820 GAAATGGAAGACATGAGAAACATGACTGACCAAGAAATGTTGACATTTGAGAGAGAG 879  
 QY 391 SerSerGlyPheSerArgGlyAlaSerIleTyraGlyValIleThrArgHisIleGlnHis 410  
 DB 880 AGCAGTGTTCTCTAGGGGGCTCTCTCACTATAGAGAGCTCACAGACATCACACCAT 939  
 QY 411 GlyArgTrpGlnAlaArgIleGlyValAlaGlyAsnLysAspLeuTyraLeuGlyThr 430  
 DB 940 GGAAGGTGGCAAGCAAGATGGTGTAGCTGCTGGAACCAAGATCTCTTACTTGGAACT 999  
 QY 431 PheSerThrGlnGlnGlnAlaGlnAlaGlnAlaTyraPheValAlaAlaIleLysPheArgGly 450  
 DB 1000 TTTGGAACCCAGAGAAAGACTCCAGAGTTACGATGACAGCAATTAATTCGCGGC 1059  
 QY 451 AlaAsnAlaValIleThrAsnPheAspIleSerArgTyraAspValGluArgIleMetAlaSer 470  
 DB 1060 ACAAAATGCTGTGACTAATCTTATATCAGAGGTACAGATGTGATCGTATATGTCTAGT 1119  
 QY 471 SerAsnLysLeuAlaGlyLysLeuAlaArgArgLysLysAspAsnAspProArgAsnLys 490  
 DB 1120 AAACACTCTTGTCTGAGAGATTAAGCGCAAGAAC----- 1155  
 QY 491 AspIleAspTyraAsnLysSerValIleThrSerValAsnAsnGlnGlnIleValGlnVal 510  
 DB 1156 -----AACACAGACATGTC-----GTACAGAAATACTGAA----- 1185  
 QY 511 GlnAlaGlyAsnAsnAsnGlnLysAspSerGluTrpLysMetValIleuPheAsnHis 530  
 DB 1185 ----- 1185  
 QY 531 ProSerGlnGlnGlnAlaAsnGlnLysAspArgGlnLysIleMetAsnCysGly 550  
 DB 1186 -----GACCAACCGCTCTAAAGCT----- 1206  
 QY 551 AsnTyraArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIleGlyIleAspSerVal 570  
 DB 1206 ----- 1206  
 QY 571 GlySerGlyGlnHisAsnMetLeuAspGlnSerLysIleGlyThrHisPheSerAsn 590  
 DB 1206 ----- 1206  
 QY 591 ThrSerSerLeuValIleThrSerLysSerSerArgGlnLysAspProGlnLysArgGly 610  
 DB 1207 -----GTTGCAAGAGGTGCTTCCACAGAAAGTCAAGTACCTCCGAGAGATC 1254  
 QY 611 ProSerLeu-----LeuPheProMetProProMetGluThrLysIleVal---AspPro 627  
 DB 1255 TTGAGTTTCCGGCGCATTTTCGGGTGCTCAAGTTATCAAAAGATGTTGGATCAAT 1314  
 QY 628 IleGlyThrSerValIleThrSerTrpLeuProSerProThrValGlnMetArgProSerPro 647

DB 1315 ATGGCGGAAATATATAGACTTGTGACATCAAAACCTATAGCTTAGCTTAAG---ACCGTC 1371  
 QY 648 AlaIleSerLeuSerHisLeuProValPheAlaSerTrpThrAspThr 663  
 DB 1372 GCTCTTACTTTGCTCCATGATGCCGATTTTCGCTGCTTGGCTGATATCT 1419  
 RESULT 13  
 US-10-425-114-11685  
 ; Sequence 11685, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Xihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(5313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 11685  
 ; LENGTH: 1231  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 701120037\_F11  
 US-10-425-114-11685  
 Alignment Scores:  
 Pred. No.: 4,06e-118 Length: 1231  
 Score: 1185.00 Matches: 237  
 Percent Similarity: 75.13% Conservative: 44  
 Best Local Similarity: 63.37% Mismatches: 63  
 Query Match: 33.84% Indels: 30  
 DB: 12 Gaps: 7  
 US-10-024-632-2 (1-663) x US-10-425-114-11685 (1-1231)  
 QY 290 AspTrpPheGlyGlnArgThrSerGlnTyraGlyValIleThrArgHisArgIlePheGly 309  
 DB 1 GATACCTTTGGCAAGAAACATCTCAATATGAGGTGTAAACAGCATAGGTGACGCT 60  
 QY 310 ArgTyraGlnAlaHisLeuTrpAspAsnSerCysLysLysGlnGlnTrpArgLysGly 329  
 DB 61 AGATATGAAAGCTCATCTATGGAGCAACAGCTGCAAGAAAGAGGCGCCAAAGAAAGAGA 120  
 QY 330 ArgGlnValTyraLeuGlnGlyTyraPheMetGlnGlnValAlaAlaGlyAlaTyraPheLeu 349  
 DB 121 AGGCAAGTTTATCTAGGGGGTTATGATGTGAAGAAAAGAGCAAGACTTATGATATG 180  
 QY 350 AlaAlaLeuLysTyraTrpGlyProSerThrHisIleAsnPheSerIleGluAsnTyraGln 369  
 DB 181 GCCGACCTCAAGTATGGGAGCCCTCTCATATTAATTTCCATTTGGAATAATATCA 240  
 QY 370 ValGlnLeuGlnGlnMetLysAsnMetSerArgGlnGlnTyraValAlaHisLeuArgArg 389  
 DB 241 AAGCACTTGAAGAAATTAAGACATACAGACAGAAAGATGTGCTCATTTAGCAAG 300  
 QY 390 LysSerSerGlyPheSerArgGlyAlaSerIleTyraGlyValIleThrArgHisIleGln 409  
 DB 301 AAAAGCAAGGATTTCTCAAGAGGAGCTTCCATGACAGAGAGTAAACAGACACCA 360  
 QY 410 HisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyraLeuGly 429  
 DB 361 CATGAGGTGGCAAGCTCGAATGGAAGGCTACCCGAAACAAAGATCTATATCTTGA 420  
 QY 430 ThrPheSerThrGlnGlnGlnAlaGlnAlaGlyValIleThrArgHisIleGlyPheArg 449  
 DB 421 ACCTTATGATCCAGAGAGAAAGACAGCTGAACCTTATGATGCTTATTAATTCAGA 480

QY 450 GYVALAASALAVATHANPHEAPPIESERATGYTASPVAGIUGRIIEMETALA 469  
 Db 481 GGGATTATCTGTCACTACTTGTATATNACANAGATGACGTGAAAAATATTGAG 540  
 QY 470 SerSerAsnLeuLeuLacIyGluLeuLaaArgArgLysLysasp---AsnAspProArg 488  
 Db 541 AGCAATACCTCTTACGACTGAACAAGCTTAACGGAAAAAGAGATGATGTGGACT 600  
 QY 489 AsnLysAspLLeaSPYrAsnLysSerValValThSerValAsnAsnGluGluThrVal 508  
 Db 601 AGAAGCCAGGCTACCGTTAACCAAAAACCTTCTACATATGACCACACTCAAGAACCAT 660  
 QY 509 GlnValGlnLacIyAsnAsnAsnGluLysAsnAspSerGluTrpLysMetValLeuPhe 528  
 Db 661 CTATGCGAG-----AAAAGTGCAGAAACCAATCATGAAATGGAAGTGTTCAGTTT 711  
 QY 529 AsnHisProSerGlnGlnGlnGlnLalaSngIyAsnGlySerAspGlnLysIleMetAsn 548  
 Db 712 CCATGCCCCCAACAGCTTGATTCAG-----NATCAAAAGATCCGAGCT 753  
 QY 549 CyseGlyAsnTyrrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIleGlyIleAsp 568  
 Db 754 TGT-----AGAACTCAGCCCTTCTCAACGCACTTATGATCAATCACTTT 795  
 QY 569 SerValGlySerGlyGlnHisAsnMetLeuAspGluSerSerLysIleGlyThrHisPhe 588  
 Db 796 -----CGTCACCAAGTGGAGAAAGGAGCAACATGGGAACATCTTG 837  
 QY 589 SerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGluLysLaseProGluLys 608  
 Db 838 TCAATCTCTTCTCTCGGTGACAAAGTTGAGTAGCTCAAGAGAAAGAACCCACGATTAAG 897  
 QY 609 ArgGlyProSerLeuLeuPheProMetProProMetGluThrLysIleValAsnProIle 628  
 Db 898 ACAAGCATGCCAACCTCTTGGAAATGCTTCAACAGTGCCTCAAAATATTG----- 948  
 QY 629 GlyThrSerValThrSerTrpLeuProSerProThrValGlnMetArgProSerProAla 648  
 Db 949 ---GCTAACGTGATCTTGGGATCTATCTTCCAAATCTCAG-----ACTGCG 993  
 QY 649 IleSerLeuSerHisLeuProValPheAlaSerTrpThrAsp 662  
 Db 994 CTTCCTAGCCTCAGATGCCAATTATTGCTTGCGTTCGACAGAT 1035  
 RESULT 14  
 US-10-424-599-123485  
 / Sequence 123485, Application US/10424599  
 / Publication No. US20040031072A1  
 / GENERAL INFORMATION:  
 / APPLICANT: La Rosa Thomas J  
 / APPLICANT: Kovalic David K  
 / APPLICANT: Zhou Yihua  
 / APPLICANT: Cao Yongwei  
 / TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
 / FILE REFERENCE: 38-21(53223)B  
 / CURRENT APPLICATION NUMBER: US/10/424.599  
 / NUMBER OF SEQ ID NOS: 285684  
 / SEQ ID NO 123485  
 / LENGTH: 1949  
 / TYPE: DNA  
 / ORGANISM: Glycine max  
 / FEATURE:  
 / OTHER INFORMATION: Clone ID: PAT\_MRT3847\_82514C.1  
 US-10-424-599-123485  
 Alignment Scores:  
 Pred. No.: 2.39e-116 Length: 1949  
 Score: 1171.50 Matches: 270  
 Percent Similarity: 56.58 Conservative: 44  
 Best Local Similarity: 48.65 Mismatches: 67

QY 562 -----GlnaspLeuIleGly 566  
 DB 968 AAGAGCGGATTCTCAGAGGGCGCTCATGTACAAAGAGTAAACAAGGCTCACAGGC 1027  
 QY 566 -----  
 DB 1028 AAGCATTTACACATGACCACTTCATGGAAGGTCAAGCTGAGATGGTAGAGTGA 1087  
 QY 567 -----IleaspSerValGlySerGly-----GlnHis----- 575  
 DB 1088 CTAAACAAGATCTA-TATCTGGCACCGGTACTCAACATGAGAGAGAGCTGAAGCTATG 1146  
 QY 575 -----  
 DB 1147 AATTGCTGCTATTAAATTACAGAGGCGAAGTCTGCTCAACTTGACATAACAGATA 1206  
 QY 575 -----  
 DB 1207 TGACGTGTAAATGACATGCGAACAACCAACCAAGTAGAGCTGTTACAGCTGCC 1266  
 QY 575 -----  
 DB 1267 CCGCCCCAGACATGATCTCAGTGAAGTGAAGTCAATAGACTCAGCCCTTCTC 1326  
 QY 576 -----AsnMetLeuaspGlySerSerIleGlyThrHis 587  
 DB 1327 AAGGACTTAGATAACATGTTTCTCAACAGTGGAGAAAGGAGCAATGGGAAACACAC 1386  
 QY 588 PhaseIleThrSerSerLeuValIleThrSerSerSerSerArgIleGlyLeuSerProGlu 607  
 DB 1387 TTGTCAATCTCTCTCTCTGAGCAAGTTAGTCAAGTCAAGAGAGAGAGCCAGAT 1446  
 QY 608 LysArgGlyProSerLeuLeuPheProMetProProMetGluThrIleValLeuPro 627  
 DB 1447 AAGCAAGCATGCCAATGCTCTTGGAAATCTTCAAGTCCCAATATATG----- 1500  
 QY 628 IleGlyThrSerValIleThrSerIleuProSerProThrValGlnMetArgProSerPro 647  
 DB 1501 -----GCTAACGCGGATCTTGGGATCTATCTCCAACTCCAG-----ACT 1542  
 QY 648 AlaleuSerLeuSerHisLeuProValPheAlaSerIleThrAsp 662  
 DB 1543 GCGCTTCTATGCTCCAGATGCCAATTTTGTCTGCTGAGACAT 1587  
 RESULT 15  
 US-10-059-911-32  
 ; Sequence 32, Application US/10059911  
 ; Publication No. US20030159180A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fischer, Robert L.  
 ; APPLICANT: Mizukami, Yukiko  
 ; APPLICANT: The Regents of the University of California  
 ; TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
 ; FILE REFERENCE: 023070-090730US  
 ; CURRENT APPLICATION NUMBER: US/10/059,911  
 ; CURRENT FILING DATE: 2002-06-17  
 ; NUMBER OF SEQ ID NOS: 42  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 32  
 ; LENGTH: 1113  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURES:  
 ; OTHER INFORMATION: Description of Artificial Sequence: truncated or  
 ; OTHER INFORMATION: mutated ANT CDNA transgene ANTND1  
 US-10-059-911-32  
 Alignment Scores:  
 Pred. No.: 3,57e-112 Length: 1113  
 Score: 1129.50 Matches: 245  
 Percent Similarity: 62.70% Conservative: 29  
 Best Local Similarity: 56.06% Mismatches: 58  
 Query Match: 32.25% Indels: 105

DB: 14 Gaps: 11  
 US-10-024-632-2 (1-663) x US-10-059-911-32 (1-1113)  
 QY 69 LeuThrValMetProLeuLysSerAspGlySerLeuCyIleLeuGlnAlaLeuLysArg 88  
 DB 1 ATGCTGTGATGGCCACTCAGATCTGATGGTCTCTTGTCTTAATGGAAGCTTCACAG 60  
 QY 89 SerGlnThrGlnValMetValProThrSerSerProLysLeuGlnAspPheLeuGly 108  
 DB 61 TCTTCTACTCGAATCAACCATCAAGATTCATCTCCAAAGTGAAGATTTCTTGGG--- 117  
 QY 109 AlaThrMetGlyThrHisGlyIleGlySerHisGluArgGly-----LeuSerLeuAsp 126  
 DB 118 -----ACCCATCACAACACACACAGTCAACAAAGAACCCATGATCTTACCTTAGAT 168  
 QY 127 SerIleTyrTyrAsnSerGlnAsnAlaGlnAlaGlnProAspArgAspLeuSerGln 146  
 DB 169 AGTTATTTCTACACACCACTCAT-----GAGCCCAACAGACTACAACTTTCMA 219  
 QY 147 ProPhe-----ArgGlnGlnGlyHisMetSerValGlnThrHisProTyrTyr 162  
 DB 220 GAGTTCCTTAGCTTCCCTCAACACAGAACCAT----- 252  
 QY 163 SerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGlnGlnIleThrHisGlu 182  
 DB 253 -----GAGGAGAACT-----AGA 267  
 QY 183 ThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlnGlyLeuLysAsnTyr 202  
 DB 268 AATTACCGGATGAC-----CCTAGTTTGAACATGGA----- 300  
 QY 203 ValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlnGlnMetAsnCys 222  
 DB 301 -----GGGCTTTTAAATGA 315  
 QY 223 GlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGlyCysGlyGluLeuGln 242  
 DB 316 GGGGATATGGGAAATTCAA-----CAG 339  
 QY 243 SerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValIleThrAlaProSer--- 261  
 DB 340 TCACTGAGCTTATCATGAGGCTCGGTGCACAAATCTGACATGCTGCTCCACAC 399  
 QY 261 ----- 261  
 DB 400 CACCAACAAACCAAAACCAACCAACCAACCAACCAACCAACCAATCTCGAAGCT 459  
 QY 262 -----GlyThrAspSerValAlaValAlaAspAlaLysArg 273  
 DB 460 CTGTGAGACAGAGCTTGGGTTGAGACGACGACAAATGCGGCTGCCAGAGAGAGG 519  
 QY 274 GlyHis-----AlaLysLeuGlyGlnGlnGlnProValIleAspGlySerIleAsp 290  
 DB 520 GACACAGAGAGATGTTGATGTTGTTGAGAAACGATGTTCTCAAGAAATCTATGAT 579  
 QY 291 ThrPheGlyGlnArgThrSerGlnTyrArgGlyValIleThrArgHisArgTyrThrGlyArg 310  
 DB 580 ACTTTTGAACAACGAACTTCTCAATACCGAGCGGTACCAAGACATAGATGAGCTGTGAG 639  
 QY 311 TyrGlnAlaHisLeuTyrAspAsnSerCysLysGlyGlnGlnThrArgLysGlyArg 330  
 DB 640 TATGAGGCTCATCTATGAGCAATGTTTCAAGAAAGAGGTCAACAGAGAAAGAGAGA 699  
 QY 331 GlnValTyrLeuGlyGlyTyrAspMetGlnGlnLysAlaAlaArgAlaTyrAspLeuAla 350  
 DB 700 CAAGTTTATCTGGAGGATTAATGATGAGAGGAAAGCTGCTCGAGCATATGATCTTCT 759  
 QY 351 AlaLeuLysTyrTyrGlyProSerThrHisIleAsnProSerIleGlnAsnTyrGlnAla 370  
 DB 760 GCACCTCAAGTACTGGGGTCCCTTACTCAACCAATTTCTTGGGAGAAATATTCAGAA 819  
 QY 371 GlnLeuGlnGluMetLysAsnMetSerArgGlnGlnIleValAlaHisLeuArgArgLys 390

```
Db      820  GAGTTGAGACATGAGACATGACTGACAGAAATATGTCACATTGAGAGAGAG 879
QY      391  SerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHis 410
Db      880  AGCAGTGGTTTCTTAAGGGTCTTCCATCTATAGAGAGTCAACAGACATCCAGCAT 939
QY      411  GlyArgTpglnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThr 430
Db      940  GGAAGTGGCAAGCAGCATGAGTGTGAGTGTGCTGAAACAAGATCTTACTTGGAACT 999
QY      431  PheSerThrGlnGlnGlnAlaAlaGlnAlaTyrAspValAlaAlaIleLysPheArgGly 450
Db      1000  TTGGAACCCAGAGAGAGCTGCAAGAGCTTACGATGTAGCAGCATTTAAGTCCGTTGGC 1059
QY      451  AlaAsnAlaValThrAsnPheAspIleSerArgTyrAspValGlnArgIle 467
Db      1060  ACAATGTGTGACTTAATTGATATATCAGAGGTACGATGTGTATCGTATC 1110
```

Search completed: March 13, 2004, 11:23:41  
Job time : 483 secs

Wed Mar 17 08:18:56 2004

us-10-024-632-2.rst

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: March 13, 2004, 00:45:08 ; Search time 2976 Seconds

(without alignments)  
6652.764 Million cell updates/sec

Title: US-10-024-632-2  
Perfect score: 3502  
Sequence: 1 MKRINSNNNTDGNHNLG.....RSPALISHLVFWASWTD 663

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlh  
-O/cgnt 1/USPTO.spool/US10024632/runat.09032004.101208.10776/app.query.fasta.1.839  
-DB=EST -OPMT=fastap -SUFFIX=rat -MINMATCH=0.1 -LOOPCH=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blcsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=exc -HASPITZ=500 -MINLEN=0 -MAXLEN=200000000  
-USBR=US10024632.QCGN.1.1.2135.@runat.09032004.101208.10776 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEOUTRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_estnum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hc1:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hc1:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vr1:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1103	31.5	752	10	AM348127
2	920	26.3	599	14	CA230632
3	911.5	26.0	1753	11	AY109146
4	910	26.0	558	14	CA232724
5	909	25.6	545	14	CD475862
6	895	25.1	769	14	CA763156
7	879	25.1	774	14	CF510900
8	855	24.4	882	14	CK267021
9	854	24.4	555	12	BJ188928
10	854	24.4	558	12	BJ178045
11	845	24.1	748	14	CF807326
12	840.5	24.0	590	12	BM094116
13	807	23.0	644	10	AM200688
14	792	22.6	500	10	AM780688
15	790.5	22.6	552	12	BU964897
16	787.5	22.5	593	13	BM307997
17	779	22.2	1160	14	CK206573
18	758.5	21.7	679	12	BJ312281
19	750.5	21.4	690	13	CA094356
20	750	21.4	720	13	BQ864461
21	746.5	21.3	631	13	CA103041
22	718	20.5	457	14	BU828154
23	715	20.4	491	14	CF244784
24	711.5	20.3	697	14	CA189006
25	710	20.3	752	14	CA238922
26	708.5	20.2	640	13	CA100375
27	708.5	20.2	523	10	BF647766
28	695.5	19.9	674	12	BF6790852
29	694.5	19.8	513	13	BQ122372
30	690.5	19.7	585	12	BM086088
31	682.5	19.5	606	13	BU825419
32	678	19.4	704	13	BQ625052
33	676	19.4	1897	11	AY103852
34	676	19.3	935	14	CK261980
35	650	18.6	907	14	CK289457
36	648	18.5	459	10	BE800260
37	647	18.5	728	14	CF035870
38	646	18.4	420	9	AJ475492
39	643.5	18.4	419	13	BU820600
40	643	18.4	694	14	CF036846
41	640.5	18.3	545	12	BI974354
42	640	18.3	641	14	CD879292
43	639.5	18.2	842	14	CF446573
44	638.5	18.2	737	14	CD832004
45	636	18.2	500	9	AI442942

# ALIGNMENTS

RESULT 1  
AM348127/c  
LOCUS  
DEFINITION  
AM348127  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

752 bp mRNA linear EST 04-OCT-2000  
Gm210001A21A2 Gm-r1021 Glycine max cDNA clone Gm-r1021-4 3', mRNA

Sequence.  
AM348127  
GI:6645837  
EST.  
Glycine max (soybean)  
Glycine max  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;



/clone="SCJFFL3C06D04"  
 /lab host="DH108"  
 /note="Organ: Base of developing inflorescence (5cm-long);  
 Vector: pSport1; Site\_1: SalI; Site\_2: NotI; An  
 undirectional cDNA library generated from (base of  
 developing inflorescence (5cm-long)). cDNA was prepared  
 from polyA+ mRNA using Superscript Plasmid System Kit  
 (Invitrogen). The double-strand cDNAs were fractionated  
 in a Sephadose Cl-2B 40cm-columns and fragments sizing  
 between 0.8 and 1.5 kb were directionally cloned into the  
 vector. Details of each source of RNA and library  
 construction can be obtained at  
 http://suest.lad.ic.unicamp.br/public"

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,08e-79 Length: 599  
 Score: 920.00 Matches: 168  
 Percent Similarity: 97.33% Conservative: 14  
 Best Local Similarity: 89.84% Mismatches: 5  
 Query Match: 26, 27% Indels: 0  
 DB: 14 Gaps: 0

US-10-024-632-2 (1-663) x CA230632 (1-599)

QY 301 GlyValThrArgHisArgTTPThrGlyArgTyrGluAlaHisLeuTTPAspAsnSerCys 320  
 Db 2 GGCGTCACAGGAGCTAGTGGACAGGAGGATAGAGCGGCACTGTGGACAAACAGCTCC 61  
 QY 321 LysLeuGluGluGlnThrArgLysGlyArgGlnValTyrLeuGlyTyrAspMetGlu 340  
 Db 62 AAGAGGAGAGGCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 121  
 QY 341 GluLysAlaAlaArgAlaTyrAspLeuAlaLeuLysTyrTTPGlyProSerThrHis 360  
 Db 122 GAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 181  
 QY 361 IleAsnHisSerIleGluLeuTyrGlnValGlnLeuGlnMetLysAsnMetSerArg 380  
 Db 182 ATCAACTTCCTCCGTCGAGAGCTACAGAGAGGAGTGAAGAGGAGGAGGAGGAGGAGG 241  
 QY 381 GlnGluTyrValAlaHisLeuArgLysSerSerGlyPheSerArgGlyAlaSerIle 400  
 Db 242 CAGAGTACGTCGCTCACTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 301  
 QY 401 TyrArgLysValThrArgHisLeuGlnHisGlyArgTTPGlnAlaArgIleGlyArgVal 420  
 Db 302 TACCGAGGAGTCAACAGGAGCAGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 361  
 QY 421 AlaGluAsnLysAspLeuTyrLeuGlyTyrPheSerThrGlnGluGlnAlaGlnAla 440  
 Db 362 TCCGAGCAAGAGCCTTCACTTGGCACTTTCAGCAGCAGGAGGAGGAGGAGGAGGAGG 421  
 QY 441 TyrAspValAlaAlaIleLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSer 460  
 Db 422 TACAGCATCGCGGAGTCAAGTTCGGGGGCTCAACCCGTCACCACTTCGACATCAGC 481  
 QY 461 ArgTyrAspValGluArgLysMetAlaSerSerAsnLeuLysGlnGluAlaArg 480  
 Db 482 CGTCAAGAGGTCGAGCAATATGAGCCAGCAACGCTGCTCGGGGAGCACTCGCCGCG 541  
 QY 481 ArgLysLysAspAsnAspPro 487  
 Db 542 CGCAGGAGGAGCAGCAGCACC 562

RESULT 3  
 AT109146  
 LOCUS AY109146 1753 bp mRNA linear HTC 17-OCT-2002  
 ACCESSION AY109146  
 VERSION AY109146.1 GI:21212582  
 KEYWORDS HTC.

## SOURCE

Zea mays  
 Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoidae; Andropogoneae; Zea.

## REFERENCE

1 (bases 1 to 1753)  
 Hainey, C.F., Dolan, M., Mao, G.H., Vogel, J.M., Whitsitt, M.S.,  
 Arthur, L.W., Hanafey, W., Morgante, M. and Tingey, S.V.  
 Maize Mapping Project/Dupont Consensus Sequences for Design of  
 Overgo Probes  
 Unpublished (2002)  
 2 (bases 1 to 1753)

## JOURNAL

Direct Submission  
 Submitted (25-APR-2002) Maize Mapping Project, University of  
 Missouri, Columbia, MO 65211, USA

## COMMENT

If you are interested in getting corresponding physical clones,  
 these are publicly available from ZmDB and may be found by BLAST  
 searching at MSU, [maizemap.org](http://maizemap.org); ZmDB, [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu); TIGR,  
[www.tigr.org](http://www.tigr.org); or NCBI, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the  
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
 Schnable, Iowa State, then clones may be requested from ZmDB:  
[www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).

## FEATURES

source

1..1753  
 Location/Qualifiers  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4577"  
 /clone\_lib="Maize Mapping Project/Dupont Consensus  
 Library"  
 /note="this sequence is part of a project of EST  
 assemblies resulting from the application of public  
 configs to seed Dupont configs; this resource was  
 assembled by Dupont as part of a collaboration for the  
 overgo addressing of BACs in conjunction with the Maize  
 Mapping Project"

## ORIGIN

Alignment Scores:  
 Pred. No.: 9.4e-78 Length: 1753  
 Score: 911.50 Matches: 205  
 Percent Similarity: 94.85% Conservative: 44  
 Best Local Similarity: 45.15% Mismatches: 88  
 Query Match: 26, 03% Indels: 117  
 DB: 11 Gaps: 11

US-10-024-632-2 (1-663) x AY109146 (1-1753)

QY 187 AspCysSerSerLeuMetProGlnMetThrGlnGlyLeuLysAsnTTPValAlaProThr 206  
 Db 42 GACATGAGCTCAAGCTTATCCCAAC-----CATGGCTCTCTCTCC 83  
 QY 207 ArgGluPheSerThrHisGlnValLeuGlnGlnGlnMetAsnCysGly----- 223  
 Db 84 CTCTCCACACACTACCAACAGGCTTCTCAAGCTTCTTACTCTCCGTAATCCT 143  
 QY 224 MetGluAsnGluArgAsnGlyValSer-----LeuGly 234  
 Db 144 CTTGAGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 203  
 QY 235 SerValGly----- 237  
 Db 204 GGCGTCGCTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 263  
 QY 238 ---CysGlyGluLeuGlnSerLeuSerLeuSerMet----- 248  
 Db 264 GTGTGGGAGAGCTGGAGAGCATCACAGCAGGAGTCTTGCGCCACTACCCGCGGCGCA 323  
 QY 249 -----SerProGlySerGlnSerSerCysValThrAlaProSerGly 262  
 Db 324 GCTGGAGCAGCGGTGAGAGACCCCGCGGAGTACC-----GTGGCGCCATGTCGTG 377



```

263 ThrAspSerValAlaValAspAlaValArgGlyValAlaValLeuGlyGln 282
378 ACCGAGGTGGCGGGGGGAGTCCGACCAAGGCG----- 410
283 ProValHisArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyrArgGlyVal 302
411 -----AGGCGGCGCCCGGAGAGCTTGGCCGAGGACATCTTACCGGTGGCGTC 461
303 ThrArgHisArgTyrPheGlyArgTyrGlyValAlaHisIleuThrAspAsnSerCysLysLys 322
462 ACCAGGACCGCGTGGACGGGAGATATGAGCGCCACTGTGGGACAACAGCTGCCCGCG 521
323 GlnGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGluLysLys 342
522 GAGGCGCAAGCCGCAAGAGCGCAAGTCTTACCTGAGAGGCTATACAGAGAGAGAG 581
343 AlaAlaArgAlaTyrAspLeuAlaAlaIleuLysTyrTrpGlyProSerThrHisIleAsn 362
582 GCGGCTGAGCTTACGACCTCGCGCGCTCAAGTACCTGGGGGCTTCAACCAACGACCAAC 641
363 PheSerIleGluValenTyrGlnValGlnLeuGlyGlyMetLysAsnMetSerArgGlnGlu 382
642 TTCAGAGGTGCCAAGTACGAGAGAGAGAGAGAGATGAAATCCATGACGCGCGCAGAG 701
383 TyrValAlaHisIleuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArg 402
702 TTCATCGCGTCTTGGCCGAGAGAGAGAGAGAGAGGCTTCTCAGAGAGCGCTTCACTACA 761
403 GlyValThrArgHisIleGlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGly 422
762 GAGATCAACAAGGATCATCAGACGCGCGGTGACAGCGAGATCGGCGAGGTGGCGGGA 821
423 AsnLysAspLeuTyrLeuGlyTyrThrPheSerThrGlnGlnValAlaGlyValTyrAsp 442
822 AACAGAGACCTGTAAGTGGGACCTTCACTAGTCAAGAGAGAGCGGGGAGGTACGAC 881
443 ValAlaAlaIleLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyr 462
882 ATCCCTCGATCAAGTTCGGGGCTCAAGCGCGTCAACCTTGAATGACATGACCGCTAC 941
463 AspValGluArgIleMet----- 468
942 GACGTGAGAGACATCTCAGACGACCTCCCGCGGCGGAGCTACGCGTCCGCGC 1001
469 -----AlaSerSerAsnLeuLeuAla 475
1002 CCGGCAAGTCCCGCTTGACTCGCTGACGCGGGAGCGGCTGCGGCATGATGCTCGGC 1061
476 GlyGluLeuAlaArgArgGlyLysAspAsnAspProArgAsnLysAspIle----- 492
1062 GGGGCTGCTGCGCTTGGCAGGCGCACCATGCGCGCGGAGAGAGACTGATGCTCTG 1121
493 -----AspTyr 494
1122 CTCGCGCTGCACTACCAAGACAGAGAGAGAGAGAGAGTCCCGGCTTCTGCTTAC 1181
495 AsnLysSerValValThrSerValAsnAsnGluTyrValGlnValGlnAlaGlyAsn 514
1182 GAGGCTTACGGCTCCGCGCGGTGAACCTGACTTCAAGATGGCACACAGAGCGGAGC 1241
515 AsnAsnAsnGluAsnAspSer-----GluTyr-----LysMetValLeuPheAsn 529
1242 AACACAAACACCGGAGCGGCTCATATGGGGCGCCACCTCTGTGACATGATGGGCGAG 1301
530 HisProSerGlnGlnGlnGlnAlaAsnGlyAsnGlySerAsp 543
1302 CAAGACAGACAGACAGACAGAGGCAACGCGTATGCCAGTAC 1343

```

RESULT 4 CA232734 558 bp mRNA linear EST 25-SEP-2003  
 LOCUS SCRUFJ3064D08.g Saccharum officinarum Fl3 Saccharum officinarum  
 DEFINITION cDNA clone SCRUFJ3064D08 5', mRNA sequence.

ACCESSION CA232734  
 VERSION CA232734.1 GI:35298534  
 KEYWORDS EST.  
 SOURCE Saccharum officinarum  
 ORGANISM Saccharum officinarum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Saccharum.  
 REFERENCE 1 (bases 1 to 558)  
 Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.  
 The libraries that made SUCSEST  
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
 CONTACT: Arruda, P  
 CENTRO de Biologia Molecular e Engenharia Genetica  
 Universidade Estadual de Campinas  
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
 Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089  
 Email: parvada@unicamp.br  
 Clone distribution: clone distribution information can be found  
 through the Brazilian Clone Collection Center (BCCC) at  
 http://www.bccc.bccp.br  
 Plate: 064 row: D column: 08  
 Seq primer: T7 Promoter Primer.  
 Location/Qualifiers  
 1..558  
 /organism="Saccharum officinarum"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4547"  
 /clone="SCRUFJ3064D08"  
 /lab\_host="DH10B"  
 /clone\_id="Saccharum officinarum Fl3"  
 /note="Organ: Base of developing inflorescence (5cm-long);  
 Vector: pSPori1; Site 1: SalI; Site 2: NotI; An  
 unidirectional cDNA library generated from (Base of  
 developing inflorescence (5cm-long)). cDNA was prepared  
 from polyA+ mRNA using Superscript Plasmid System Kit  
 (Invitrogen). The double-strand cDNAs were fractionated  
 in a Sephadose CL-2B 40cm-columns and fragments sizing  
 between 0.8 and 1.5 kb were directionally cloned into the  
 vector. Details of each source of RNA and library  
 construction can be obtained at  
 http://succest.lad.ic.unicamp.br/public"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.75e-78 Length: 558  
 Score: 910.00 Matches: 167  
 Percent Similarity: 97.30% Conservative: 13  
 Best Local Similarity: 90.27% Mismatches: 5  
 Query Match: 25.99% Indels: 0  
 DB: 14 Gaps: 0

US-10-024-632-2 (1-663) x CA232734 (1-558)

```

301 GlyValThrArgHisArgTyrPheGlyValAlaHisIleuThrAspAsnSerCys 320
2 GCGGCTACCAAGACATGATGAGACAGAGAGATGAGCGCACCTGTGGGACAACAGCTGC 61
321 LysLysGlnGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlu 340
62 AAGAGAGAGAGCGACACAGACAGAGAGGCGAGGCAAGTTATCTCGCGCGGTATGACATGGAG 121
341 GlnLysAlaAlaArgAlaTyrAspLeuAlaIleuLysTyrTrpGlyProSerThrHis 360
122 GAGAGAGGCGGCGAGGCTTACGACCTGGCGGCTTCAAGTCTGGGCGCTTCAACCAAC 181
361 IleAsnPheSerIleGluAsnTyrGlnValGlnLeuGlyGlyMetLysAsnMetSerArg 380
182 ATCAACTTCCGCGTGGAGACTACCAAGAGAGAGCTGAGAGATGAGAGACATGACACGG 241
381 GlnGluTyrValAlaHisIleuArgArgLysSerSerGlyPheSerArgGlyAlaSerIle 400

```





/clone\_11b="Vitis vinifera cv. cabernet sauvignon (Clone 8) Bud - CABUD"  
 //note="Organ: Bud; Vector: pDNR; Site 1: SfiI; Site 2: SfiI; CABUD is a cDNA library of Vitis vinifera cv. 'Cabernet Sauvignon', Clone 8 dissected buds. Samples were collected May 13, 2002 from pre-bloom plants (10-11 days before bloom), pre-veraison. Sampled vines were located at the University of California, Davis, Experimental vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:  
 5'-AAGCAGGTGATTCACGACGAGTTCGACCGCGG-3' and 5'-ATTCTAGAGCCGAGCGCGGACATG-3' (30) NM-3' library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 Kb size fraction."

## ORIGIN

## Alignment Scores:

Pred. No.:	Length:	774
Score:	879.00	Matches: 176
Percent Similarity:	79.23%	Conservative: 30
Best Local Similarity:	67.68%	Mismatches: 34
Query Match:	25.10%	Indels: 20
DB:	14	Gaps: 8

US-10-024-632-2 (1-663) x CFS10900 (1-774)

QY 133 GlnanalaGluAlaGlnProAsnArg-----AspLeuLeuSerGlnProPhe 148  
 Db 774 CAATATCGGACCGACCTGACCAACATTCCTTCAACCTTCTTCAAGATCTTT 715  
 QY 149 ArgGlnGlnGly-----HisMetSerValGlnThrHisProTyrTyrSerGlyLeuAla 166  
 Db 714 AGGCAACAGAGTCCAGCAACCAATTCACCTTCAACCAACCTTCACTGACCTGGCC 655  
 QY 167 CysHisGlyLeuTyrGlnAlaProLeuGluGluGluTyrThrTyrLeuGluThrHisValSer 186  
 Db 654 TCGATGGATTTACCAACCCATTTAAGAAAGAA--ACAAAGAAAGCCCACTTGA 558  
 QY 187 AspCysSerSerLeuMetProGlnMetThrGluGly-----LeuLysAsnTyr 202  
 Db 597 GATTCCTCAATCCCAATCCCTCTCTATGGGAGAGATGGAATGCTTCCGAAAACTGG 538  
 QY 203 ValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGluGlnGlnMetLeuAsnCys 222  
 Db 537 GTTGA-----AGGCACTACTCGACTAC--CATGCATAGAGCGACGATGACAC 487  
 QY 223 GlyMetGlyAsnGluArgAsnGlyVal--SerLeuGlySerValGlyCysGlyGlyLeu 241  
 Db 486 AGCATCGTTGATGACGAGAGAGCTCTCGGTCTATTTGAGCTATGAGGCTTGTGGGATTTA 427  
 QY 242 GlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValThrAla----- 259  
 Db 426 CAGTCTTCAAGCTTCTCAAGCCCGCTTCTAGTCAAGCTTGTGAAGCTTCCACAGA 367  
 QY 260 -----ProSerGlyThrAspSerValAlaValAspAlaLysLysAsnGlyHisAla 276  
 Db 366 CAGATCTCGCTACCGGAGAGCTGAGTGTGCTGCTGCAACAAAGAAAGAGGAGTCCGGA 307  
 QY 277 LysLeuGlyLysLysGlnProValHisArgLysSerLeuAspThrPheGlnGlnGlnGly 286  
 Db 306 AAGGTTCGCAAAAGCAACCACTTCAAGAGTCAATTCATGACATTTGGCGAGAGACT 247  
 QY 297 SerGlnTyrArgGlyValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTyr 316  
 Db 246 TCACATATAGAGGAGTTACAGGATGATGATGATGATGATGATGATGATGATGATGAT 187  
 QY 317 AspAsnSerCysLeuLysGlyGluGlnThrArgLysGlyArgGlnValTyrLeuGlyGly 336  
 Db 186 GATTAATAGTTCGAAGAAGAGGAGGAGCTAGAGAAAGAGACAAAGTTATCTGGGGGG 127  
 QY 337 TyrAspMetGluGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrGly 356

Db 126 TATGATATGAGAGAGAAAGCTGCAAGAGCTTACGATCTGGCGCCCTCAATATCTGGGCA 67  
 QY 357 ProSerThrHisLeuAsnPheSerLeuGluAsnTyrGlnValGlnLeuGluGluMetLys 376  
 Db 66 CTTTCTACTCATATCACTTCCCTTGAATAATTCACCAAAAAAAAAAAAAAAAAAAAAA 7

RESULT 8  
 CK267021 882 bp mRNA linear EST 12-DEC-2003  
 LOCUS EST13099 potato abiotic stress cDNA library Solanum tuberosum cDNA  
 DEFINITION clone POAC822 5' end, mRNA sequence.  
 ACCESSION CK267021 GI:39823999  
 VERSION CK267021.1 GI:39823999  
 KEYWORDS EST.  
 SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Solanales; Solanaceae; Solanum.  
 1 (bases 1 to 882)  
 Buell, C.R., Hart, A., Zismann, V., Karamychava, S.A. and Baker, B.  
 Generation of ESTs from abiotic stressed potato tissue  
 Unpublished (2003)  
 Other ESTs: EST13100  
 Contact: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from TIGR via potato@tigr.org  
 Seq primer: ATT TAG GTG ACA CTA TAG.  
 Location/Qualifiers  
 1..882  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultiivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POAC822"  
 /lab\_host="DH10B-Toma"  
 /tissue\_type="abiotic stress treated leaf and root tissue"  
 /clone\_11b="potato abiotic stress cDNA library"  
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample.."

## ORIGIN

## Alignment Scores:

Pred. No.:	Length:	882
Score:	855.00	Matches: 171
Percent Similarity:	70.63%	Conservative: 31
Best Local Similarity:	59.79%	Mismatches: 58
Query Match:	24.41%	Indels: 26
DB:	14	Gaps: 4



QY	34	GLIUYVAIAIAAAGAAATYRASPLeuAlaIAleuysTYRTPGIYP:oseThRHis	360
Db	190	GACANGCAGCCAAAGCCTTAGACTTGGCAACCGCTCAAAATTTGGGGTGCAGCACCAACC	249
QY	361	IIeAnPheSerIleGIUAenTYrGLIYAIdInLeuGIUWetLYPAnMeCSeRArg	380
Db	250	ATCAACCTCAGCGTTGATACATACGACAAAGAGCGTGAGGAGAGAAAGAAATCTTCGG	309
QY	381	GIInGIUTYrVALIAhIEleuArgRPySseSerGIYPheSerATSGIYALASerIle	400
Db	310	CAGAGATATGTGGCTCGTTAGAAAGGAAAGCGATGGCTTTTCAGAGGAGCATCCATG	369
QY	401	TYRArgIYVALThnRGHISHeGIInHISGIYATGTTPGIInIAArgIleGIYArgVAL	420
Db	370	TATGAGGGGCGTGCAGAGACCACTACGATGAGATGGAGATGGCAAGCAGCATTTGGCGGAGTT	429
QY	421	AlaGIYAenLYPASPLeuTYrLeuGIYThnRPheserThrGIInGIUInALAIGLIUALA	440
Db	430	GCGGGAAACAAGAGACCTTACCTGGGCACTTAGCAGCACTCAAGAGGAAGCGCGAGGCC	489
QY	441	TYrAspVALIAIAIAIleIysPheRgGLIYALASnAlaVALThnAsnPheAspIleSer	460
Db	490	TACGACATAGCTGCATAAATATGCTGGAATTAATGTCGGTACAAACTTTCAATATCTCT	549
QY	461	ArgTYr 462	
Db	550	CGTTAC 555	

RESULT	10
BJ178045	
LOCUS	
DEFINITION	Bj178045 558 bp mRNA linear EST 16-OCT-2003 caulonemata and malformed buds Physcomitrella patens subsp. patens
ACCESSION	Bj178045
VERSION	Bj178045.1 GI:18346002
KEYWORDS	EST.
SOURCE	Physcomitrella patens subsp. patens
ORGANISM	Physcomitrella patens subsp. patens Embryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella. 1 (bases 1 to 558) Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Shimozaki,K., Kohara,Y. and Hasebe,N. Comparative genomics of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
JOURNAL	

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)  
22709184  
12808149

Contact: Tadao Shin-i  
Center for Genetic Resource Information  
National Institute of Genetics  
1111 Yaba, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6855  
Fax: 81-559-81-6855  
Email: tshinigenes.nig.ac.jp

A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-F1C phage vector (Camici et al 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'- gagagagagagagatccacacctgagagagttttttttttttttttn-3' was used as a 1st 3' primer, and 5'- gatttcctagtcgcatcgctggttcacagacgagatgactcgagacacggnnnnn-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, agggccaaatccggccagctccgattctgctgagacccg). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).

Probenamata were blended by the POLYTRON, and then cultivated on the BCDNA6 medium for 13- 14 days under the continuous light.

FEATURES  
SOURCE

These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PHYSCObase (<http://mosa.nibb.ac.jp/>).

ORIGIN

Alignment Scores:	
Pred. No.:	5_41e-73
Score:	854.00
Percent Similarity:	93.4%
Best Local Similarity:	66.81%
Query Match:	29.3%
DB:	12
Gaps:	0
Length:	5
Matches:	5
Conservative:	1
Mismatches:	1
Indels:	0
Gaps:	0

US-10-024-632-2 (1-663) X BU178045 (1-558)

Qy	281	lysglnprovalhlsaglylsserileasphrheglglnatgtnserglntyrarg	300
Db	10	AAATATCCGCCCCCGCAAGTCATGATACCTCCGCAAGAAACATCTGTGACGA	69
Qy	301	gllyvalthraarghlsargttrpthglyargtyglualahlsleutrpasphaserlys	320
Db	70	GGCTGCACTAGGACCTGTTGGACTGGACGGATACAGGACACATCTGTGGAAACAACTTGT	129
Qy	321	lyelysaglucglvgnthtrarglysglyargglinaltyrleuglygltytraspmetglu	340
Db	130	AGAAAGAAAGGTCAAATCTCGCAAGTCGACAACTATACCTTAGAGAGATATGATTAAGAA	189
Qy	341	glulysalaaiaargalatyraspleualahleulyslyrtrpglproserthrls	360
Db	190	GACAAAGGACGACCAAGCCTATGACTTGGACGCGCTCAAAATTTGGGTGTCAGACCCACC	249
Qy	361	lleasnpeserilegluasntyrglinalglucugluglmetclysasmetsezararg	380
Db	250	ATCAATCTCAGCCTTGATACATACGACAAAGCTGGAGGAGATGAAGAATGTCCTGG	309
Qy	381	glnqlutyralahlsleulazarglylsergerglyphaserarglylaserile	400
Db	310	CAGAGATATGTGGCTCGTTGAGAAAGAAAGCAGTGGCTTTTCAGAGGAGACATCCATG	369
Qy	401	tyrarglyvalthtrarghshlsglnhslgylargttrpglinalaarglleglyarval	420
Db	370	TATCAGAGGGGTGACGAGACACCATCAGATGGAGATGGCAAGCAGCATTTGGCGAGTT	429
Qy	421	Ala glyasnlysasp leu tyrleu gly thrphaserthrnglnuglnualahleulua	440
Db	420	GCGGAAACAAGACCTGTACTGTGGCACTTACAGCACTCAAGAGAAAGCGGAGGGCC	489
Qy	441	tyraspvalaaialahlelysphearglylalaasnalaalathrasnpheaspi leser	460
Db	430	TACGACATAGCTGCCATAAATATCTGTGGAATTATGCGGTCAAACTTTCATATCTCT	549
Qy	461	Arg tyr 462	
Db	550	CGTTAC 555	

RESULT 11

CF807326

LOCUS

DEFINITION

peh8023xf07 USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation Phytophthora sojae cDNA clone SH8023f07 5, mRNA sequence

748 bp

mRNA

linear

EST 27-OCT-2001



ACCESSION CR807326  
 VERSION CR807326.1 GI:37995737  
 KEYWORDS EST.  
 SOURCE Phytophthora sojae  
 ORGANISM Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.  
 1 (bases 1 to 748)  
 REFERENCE Tyler B.  
 TITLE Tyler B. Not Published  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Tyler B  
 Tyler Lab  
 VBI  
 1880 Pratt Dr., Blacksburg, VA 24061, USA  
 Tel: 540-231-7318  
 Email: bmttyler@vt.edu  
 PCR primers  
 FORWARD: BK reverse primer  
 BACKWARD: BK reverse primer  
 Plate: 023 row: F column: 07  
 Seq primer: BK reverse primer  
 High quality sequence stop: 748.  
 Location/Qualifiers  
 1..748  
 /organism="Phytophthora sojae"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:67593"  
 /clone="SHB023F07"  
 /tissue\_type="mycelium"  
 /cell\_line="P6497"  
 /dev\_stage="48 hr. post infection stage"  
 /lab\_host="Soybean plant"  
 /clone\_lib="USDA-IFRS:Expression of Phytophthora sojae genes during infection and propagation"  
 /note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"  
 ORIGIN  
 Alignment Scores: 6,92e-72 Length: 748  
 Pred. No.: 845.00 Matches: 154  
 Score: 92.90% Conservative: 16  
 Percent Similarity: 84.15% Mismatches: 13  
 Best Local Similarity: 24.13% Indels: 0  
 Query Match: 14 Gaps: 0  
 DB: 14  
 US-10-024-632-2 (1-663) x CR807326 (1-748)  
 QY 281 TysGlnPrcvAlHlsArgLysSerlleApHrPhiegLyGlnArgThSerGlnTyArg 300  
 Db 198 CAGAAACCTCTCCAAAGAAAACCGCGACCTTCGGCAACGACCTCCATCTACGCG 257  
 QY 301 GtYvAlThrArgHlsArgTrpHrGtArgTyrGtAlHlsLeuTrpAspAsnSerCys 320  
 Db 258 GGGGTACCCGAGATGATGAGCGGAAATTCGAAGCTCATCTTGGGCAATGTTGT 317  
 QY 321 TysLysGlnGlyGlnThrArgLysGtArgGlnValTyrLeuGlyGlyTyrAspMetLu 340  
 Db 318 AGAAGAGAAAGGCAAGCAAGAAAGAAAGAAAGATTACCTGGGTGATTATGACAAAGAA 377  
 QY 341 GtLysAlAlAlArgAlaTyrAspLeuAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 360  
 Db 378 GAAAGAGCAAGCGGCTTACGATCTCGGAGCTCTAACTCTGGGCTCACTGCAAC 437  
 QY 361 lLeAsnPhseerlleGluSerTyrGlnValGlnLeuGluGluMetLysAsnVetSerArg 380  
 Db 438 ACCAATCTTCCATTCACATCTATGAGAGAACTGAGAGATGAGAAACATGACCAAG 497  
 QY 381 GtGlnGtYvAl 400  
 Db 498 CAAAGAGTTGTCTCTCTACCAAGAAAGCAAGGAGTGTCTTGGGGGCGCTCTATA 557  
 QY 401 TysArgGtYvAlThrArgHlsArgHlsGlnHlsGlyArgTrpGlnAlAlArglleGtYvArg 420

Db 558 TACAGAGAGTGTACAGACACACCATGCGCCGAGCGGAGCAATTCGACAGCTT 617  
 QY 421 AlAgIyAsnLysAspLeuTyrLeuGlyThrPhseerThringLingLingLingLingLing 440  
 Db 618 GCCGGAACCAAGACCTTACCTTGGAACTTTCAGACACCAAGAAAGAGCTGCGAGGCC 677  
 QY 441 TysAspValAl 460  
 Db 678 TATGACATTGCTGCTATCAAAATTCAGGGGATTAAATGATGATCAGCACTTTCATGAGT 737  
 QY 461 ArgTyrAsp 463  
 Db 738 CCTACGAT 746  
 RESULT 12  
 BM094116  
 LOCUS  
 DEFINITION  
 accession  
 version  
 key words  
 source  
 organism  
 EST.  
 Glycine max (soybean)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 1 (bases 1 to 590)  
 Shoemaker R., Kerm P., Vodkin L., Expelding J., Corryell V., Khanna A., Bolla B., Marra M., Hillier L., Kucaba T., Martin J., Beck C., Wylie T., Underwood K., Steptoe M., Theising B., Allen M., Bowers Y., Person B., Swaller T., Gibbons M., Pape D., Harvey N., Schurk R., Ritter E., Kohn S., Shin T., Jackson Y., Cardenas M., McCann R., Waterson R. and Wilson R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 Contact: Shoemaker R./Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: cu@resgen.com  
 High quality sequence stop: 431.  
 Location/Qualifiers  
 1..590  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl036-2554"  
 /tissue\_type="somatic embryos cultured on MSD 20"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-cl036"  
 /note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed in the laboratory of Dr. Ila Vodkin by Anu Khanna at the University of Illinois at



ORIGIN

Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

Alignment Scores:

Pred. No.:	1,26e-71	Length:	59
Score:	840.50	Matches:	150
Percent Similarity:	90.77%	Conservative:	21
Best Local Similarity:	80.00%	Mismatches:	17
Query Match:	24.00%	Indels:	1
DB:	12	Gaps:	1

US-10-024-632-2 (1-663) x BM094116 (1-590)

[illegible]

RESULT 13	644 bp	mRNA	linear	EST 02-DEC-2001
AW200688				
LOCUS				
DEFINITION	644 bp	mRNA	linear	EST 02-DEC-2001
	max cDNA clone	GENOME SYSTEMS CLONE ID		
	se92c07.y1 Gm-cl027	Glycine max		

TITLE  
JOURNAL  
COMMENT

Schurr, R., Ritzer, E., Kohn, S., Shint, T., Jackson, Y., Cardenas, M.,  
 McCann, R., Waterson, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: [est@wsonu.wustl.edu](mailto:est@wsonu.wustl.edu)  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35891 For further information  
 call: (800)-533-4363 or contact via email: [cu@resgen.com](mailto:cu@resgen.com)  
 Insert Length: 968 Std Error: 0.00  
 High quality sequence stop: 418.

**FEATURES**  
**source**

1. .644  
/organism="Glycine max"  
/mol\_type="msna"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl027-229"  
/tissue\_type="cotyledons of 3- and 7-day-old Williams seedlings"  
/lab\_host="DH10B"  
/clone\_host="SK-c1027"  
/note="Vector: pBluescript II SK+; Site1: EcoRI; Site2: XhoI; This cDNA library was constructed from mRNA isolated from cotyledons of 3- and 7-day-old Williams seedlings which were propagated on paper towels with distilled water. The cotyledons were flash-frozen in liquid nitrogen, then lyophilized for 72 hours. Unequal amounts of mRNA was used for cDNA synthesis. Stratagene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First- stranded synthesis was performed with 5'-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An anchor nucleotide (5'-A, C, or G) was added to the 3' end of the primer [GAGAGGAGAGGAGAGGAGACTGCTCCGAG(T)18] to anchor the primer at the 5' end of the poly(A) tract. After second- strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using Glycerol Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript (tm) II XR Predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 97% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (m30). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

ORIGIN

Alignment Scores:	
Pred. No.:	2.82e-68
Score:	807.00
Percent Similarity:	92.61%
Best Local Similarity:	84.09%
Query Match:	23.04%
DB:	10
	0
Length:	644
Matches:	18
Conservative:	15
Mismatches:	13
Indels:	0
Gaps:	0

US-10-024-632-2 (1-663) X AW200688 (1-644)

QY 281 LysGlnProValHisArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyrArg 300  
 Db 114 CAGAAACCCCTCTCCAAAGAAAACCGTCGACACTTCGGCCCAAGCACTCCATCTACCGC 173

Oy	301	GLVLAITPRASGHAISATGTPRTHGIIATGTYGGLAAIAHLSLENTPRASPSNERYS	320
Db	174	GGCGGACCCGACATAGATGGACGGGAAGATCGAAGCTCATCTATGGGACAATAGTGGT	233
Oy	321	LYSLVSGIUGIYGLINTHRATGLVSGIYARGIIVALTYTLEUGIYGLITYTPASMEGLU	340
Db	234	AGAGAGAGAGGCGCAAGACGAGAAAGAAACAAAGTTTCCGCGGTGATTATGACAAAGAA	293
Oy	341	GLIYVSAIAIAATGATATATYTPASPLEUAAIALEULVSYTYTPGIVPROSETHRHS	360
Db	294	GATTAAGGACGACGAGGCTTACATGTCGCACTTCGCAAGTACTGGGGTCCACTACACC	353
Oy	361	ILEAPHPSEETLEGLUASNTYTGIVVALGINTLEUGIUGIUMETLYASMESEERARG	380
Db	354	ACCAACTTTCATTTCCACTATGAGAAAGAACTGAGAGATGAAACATGACCAAG	413
Oy	381	GLNGIUTYVALAHISLEUATGRHGRJYSSERSGIYPHESERARGIYALASERTLE	400
Db	414	CAAGGATTGTGGCTCTCTACGAAGAAAGACATGGTTTCTTAAGGGGGGCTCTATA	473
Oy	401	TYTATGIVYALTHTRARGHISHISGLINHISGIYATGTPGINALATGILEGIYARGAL	420
Db	474	TACAGAGAGATGACAGACACCAACGACATGGCCGATGGCAGCGAGAAATGGCAGAGTT	533
Oy	421	ALAGIYASNLVSPLEUTYTRLEUGIYTHRPHESERTHRINGIUGIUNIAALAGIUNIA	440
Db	534	GCGGAAACCAAGACCTCTACCTTGGAATTCACACACCCAAAGAAAGCTGCTGANGCC	593
Oy	441	TYTASPVVALAIAIILEYSPHEARGGIYALASNDALAVALTTHASN	456
Db	554	TATGACATGTCCTGCTATCAAAATTCAGCGCATTAATATGACATGACAAAC	641

Db 363 CAGCGAGAAATGACGACAGTTGCGGAAACAAAGACCTTACTTGAACCTTCACACC 422

Qy 434 GINGLUJUALAALAGLUALATYRASPVALAALALALEYSPHEARGLYALAENALIA 453

Db 423 CAAGAGAGAGCTGTGAGGCTATGACATTCGTCTATCAAAATTCAGGGATTAATGCA 482

Qy 454 ValThrasphenheapile 459

Db 483 GTCACAAACTTGACATG 500

RESULT 15

BU964897 552 bp mRNA linear EST 21-OCT-2002

LOCUS BU964897

DEFINITION Bat04807 y1 Gm-cl036 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl036-12685 5' similar to TR:Q9SYC2 Q9SYC2 FLIM15.6 PROTEIN. ; mRNA sequence.

ACCESSION BU964897

VERSION BU964897.1 GI:24205644

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max (soybean)

REFERENCE 1 (bases 1 to 552)

AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Epeiding, U., Coryell, V., Khanna, A., Bolla, B., Marras, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bower, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schuur, R., Ritzer, B., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE Public Soybean EST Project

JOURNAL Unpublished (1999)

COMMENT Contact: Shoemaker R./Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Reggen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@reggen.com web site: www.reggen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 445.  
Location/Qualifiers  
1..552  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-cl036-12685"  
/tissue\_type="somatic embryos cultured on MSD 20"  
/lab\_host="DH10B"  
/clone\_id="Gm-cl036"  
/note="Vector: pSPORT1, Site 1: NotI, Site 2: SalI. This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies Superscript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

FEATURES

Source

ORIGIN

Alignment Scores:

Score:	8.91e-67	Length:	552
Percent Similarity:	790.50	Matches:	146
Best Local Similarity:	51.21%	Conservative:	20
Query Match:	80.22%	Mismatches:	15
DB:	22.57%	Indels:	1
	13	Gaps:	1

US-10-024-632-2 (1-663) x BU964897 (1-552)

Qy 301 GlyValThrGhiHisArgTyrThrGlyValGlyValHisLeuTyrAspAsnSerGly 320

Db 3 GGAGTAAAGCGGACATGATGAGCTGGAAGGATGAGTCACTTGGGAAATGAGCTGT 62

Qy 321 LysLysGluGluGluThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlu 340

Db 63 AGAGGGAGAGGCGCAATCAAGAAAGAGCCCAAGTTTGGGTGATATGATTAAGAA 122

Qy 341 GluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyProSerThrHis 360

Db 123 GAAAAAGCAGCTAGGCTTATGATTTACCTGACCTGAAGTACTGGGGGACATCCACACT 182

Qy 361 IleAsnProSerIleGluAsnTyrGlnValGlnLeuGluGluMetLysAsnMetSerArg 380

Db 183 ACCAACTTCCAAATTAAGTAACTATGAGAAAGAAATGATGAATTAACACATGCGGA 242

Qy 381 GINGLUYVALAALAHISLEUARGARGLYSSERSERGLYPHESERARGLYALASERILE 400

Db 243 CAAGATTGTGCTGCGCCATAGAGAAAGACAGTGTCTCCAGGGGTGCATCAATG 302

Qy 401 TyrArgGlyValThrArgHisIleGlnHisGlyArgTyrGlnAlaArgIleGlyArgVal 420

Db 303 TATCGAGAGTTACCAAGCATCAACACACGAGAAATGGCAAGAAATGGCAGATT 362

Qy 421 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheSerThrGlnGluAlaAlaGluAla 440

Db 363 GAGAGAAACAAAGATCTTACTTGGGACTTTCATCTAGAAAGAGGCTGCTAAGCA 422

Qy 441 TyrAspValAlaAlaAlaLysPheArgGlyAlaAlaAlaValThrAspPheAspIleSer 460

Db 423 TAGCATATGCTGCGATTAAGTTCAAGGCTCAACAGCTGTCACAAACTTGCATGAGC 482

Qy 461 ArgTyrAspValGluArgGlyLeuAlaSerSerAsnLeu---LeuAlaGlyGluLeuAla 479

Db 483 CGCTAGACGTGAAGCCATCTTGAAGCAACACTTCCCAATGAGAGAGACTGCA 542

Qy 480 ArgArg 481

Db 543 AAGCGT 548

Search completed: March 13, 2004, 07:59:35

Job time : 2990 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2004, 10:47:09 ; Search time 38 Seconds

(without alignments)  
3684.069 Million cell updates/sec

Title: US-10-024-632-2

Perfect score: 3502

Sequence: 1 MKRINSSNNDDGNNHNLG.....RSPPAISLHLPVFASTWDT 663

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 809742 seqs, 21153259 residues

809742

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCIT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCITUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3502	100.0	663	US-10-024-632-2	Sequence 2, Appl
2	1715	49.0	665	US-10-024-632-4	Sequence 4, Appl
3	1408.5	40.2	642	US-10-024-632-6	Sequence 6, Appl
4	1349.5	38.5	669	US-10-024-632-9	Sequence 9, Appl
5	1338	37.9	555	US-10-059-911-2	Sequence 2, Appl
6	1319	37.7	555	US-10-059-911-21	Sequence 21, Appl
7	1313	37.5	555	US-10-059-911-22	Sequence 22, Appl
8	1307	37.3	548	US-10-059-911-5	Sequence 5, Appl
9	1233.5	35.2	452	US-10-059-911-13	Sequence 13, Appl
10	1234	35.0	473	US-10-059-911-9	Sequence 9, Appl
11	1180.5	33.7	585	US-10-024-632-11	Sequence 11, Appl
12	1129.5	32.3	370	US-10-059-911-15	Sequence 15, Appl
13	1042	29.8	308	US-10-059-911-10	Sequence 10, Appl
14	955.5	27.3	277	US-10-059-911-11	Sequence 11, Appl
15	948	27.1	574	US-10-374-780A-334	Sequence 334, Appl

16	947.5	27.1	205	US-10-059-911-17	Sequence 17, Appl
17	939.5	26.8	275	US-10-059-911-25	Sequence 25, Appl
18	939.5	26.8	415	US-10-374-780A-400	Sequence 400, Appl
19	927	26.5	558	US-10-374-780A-2496	Sequence 2496, Appl
20	891.5	25.5	498	US-10-374-780A-346	Sequence 346, Appl
21	889	25.4	314	US-10-059-911-23	Sequence 23, Appl
22	861	24.6	174	US-10-059-911-19	Sequence 19, Appl
23	762.5	21.8	357	US-10-059-911-14	Sequence 14, Appl
24	731.5	20.9	490	US-10-374-780A-1765	Sequence 1765, Appl
25	682	19.5	430	US-09-934-455-154	Sequence 154, Appl
26	682	19.5	430	US-10-286-264-104	Sequence 104, Appl
27	682	19.5	430	US-10-094-458A-6	Sequence 6, Appl
28	682	19.5	430	US-10-295-403-34	Sequence 34, Appl
29	682	19.5	430	US-10-374-780A-202	Sequence 202, Appl
30	681	19.4	430	US-10-094-458A-3	Sequence 3, Appl
31	675.5	19.3	443	US-10-374-780A-1311	Sequence 1311, Appl
32	666.5	19.0	313	US-10-225-067-140	Sequence 140, Appl
33	666.5	19.0	313	US-10-374-780A-2164	Sequence 2164, Appl
34	658.5	18.8	275	US-10-059-911-16	Sequence 16, Appl
35	638	18.2	415	US-10-374-780A-1314	Sequence 1314, Appl
36	633	18.1	399	US-10-374-780A-1313	Sequence 1313, Appl
37	614	17.5	622	US-10-374-780A-1743	Sequence 1743, Appl
38	611.5	17.5	255	US-10-024-632-13	Sequence 13, Appl
39	567	16.2	429	US-10-374-780A-1312	Sequence 1312, Appl
40	566	16.2	426	US-10-374-780A-1310	Sequence 1310, Appl
41	558.5	15.9	196	US-10-059-911-12	Sequence 12, Appl
42	532	15.2	264	US-10-059-911-24	Sequence 24, Appl
43	517	14.8	205	US-10-094-458A-8	Sequence 8, Appl
44	483.5	13.8	428	US-09-934-455-42	Sequence 42, Appl
45	483.5	13.8	428	US-10-374-780A-536-82	Sequence 82, Appl

#### ALIGNMENTS

RESULT 1  
US-10-024-632-2  
Sequence 2, Application US/10024632  
Publication No. US20020170093A1  
GENERAL INFORMATION:  
APPLICANT: Monsanto Technology LLC  
APPLICANT: He, Steve S.  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION  
TITLE OF INVENTION: GROWTH AND USES THEREOF  
FILE REFERENCE: 38-21(51837)B  
CURRENT FILING DATE: 2001-12-19  
PRIOR FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: US 60/257,896  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 663  
TYPE: PRT  
ORGANISM: Glycine max  
US-10-024-632-2

Query Match  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKRINSSNNDDGNNHNLGSPHMKMEATSAATVPTTFYNSPSQSHLSNFGMCYGVG 60
DB	1	MKRINSSNNDDGNNHNLGSPHMKMEATSAATVPTTFYNSPSQSHLSNFGMCYGVG 60
QY	61	ENGHFSPLTWPKSPSGSLCEALRSQTOVWVPTSPKLBDFLGATMGREHGSHE 120
DB	61	ENGHFSPLTWPKSPSGSLCEALRSQTOVWVPTSPKLBDFLGATMGREHGSHE 120
QY	121	RGLSLDSIYYNSQVAEAQPNRDLISQPPROQGHMSVOTHPYVGLACHGLYQAPLEERT 180
DB	121	RGLSLDSIYYNSQVAEAQPNRDLISQPPROQGHMSVOTHPYVGLACHGLYQAPLEERT 180

QY 181 KETHVSDSSIMPQMTBGLKNVAVPTREFTTHQVLEEQNCKGNRNGVLSGVCCG 240  
DB 181 KETHVSDSSIMPQMTBGLKNVAVPTREFTTHQVLEEQNCKGNRNGVLSGVCCG 240  
QY 241 LOSLSLSMPSGSSCVTAPSGTDSVAADAKRGHAKLGQKQPVHRKSIDTFGQRTSQR 300  
DB 241 LOSLSLSMPSGSSCVTAPSGTDSVAADAKRGHAKLGQKQPVHRKSIDTFGQRTSQR 300  
QY 301 GVTFRHRTGRYEALHMDNSCKKEGTRKGRQVYIGGYDMEKARAYDLALAKTWGPSTH 360  
DB 301 GVTFRHRTGRYEALHMDNSCKKEGTRKGRQVYIGGYDMEKARAYDLALAKTWGPSTH 360  
QY 361 INFSTENYQVLEEMKMSRQEVYAHLRKSSGFSRGASIRGVTRHHQGRWQARIGRV 420  
DB 361 INFSTENYQVLEEMKMSRQEVYAHLRKSSGFSRGASIRGVTRHHQGRWQARIGRV 420  
QY 421 AGNKDLYLGTSTOEEAAEAYDVAIKERGANAVTNPDISRYDVERIMASSNLLAGELAR 480  
DB 421 AGNKDLYLGTSTOEEAAEAYDVAIKERGANAVTNPDISRYDVERIMASSNLLAGELAR 480  
QY 481 RKXNDPBNKIDYKRSVYTVNNEETVOYAGNNDNDEEMKQVLFNHPSCQOQANGN 540  
DB 481 RKXNDPBNKIDYKRSVYTVNNEETVOYAGNNDNDEEMKQVLFNHPSCQOQANGN 540  
QY 541 GSDOKIMQGNYSNRSFAMALQDLIGIDVSGQHMLDESSKIGTFHFSNTSLVTSLS 600  
DB 541 GSDOKIMQGNYSNRSFAMALQDLIGIDVSGQHMLDESSKIGTFHFSNTSLVTSLS 600  
QY 601 SEASPEKRGPSLLIPMPMETKINPIGTSTSLPPTQMPSPALISHLPVFASW 660  
DB 601 SEASPEKRGPSLLIPMPMETKINPIGTSTSLPPTQMPSPALISHLPVFASW 660  
QY 661 TDT 663  
DB 661 TDT 663

## RESULT 2

US-10-024-632-4  
; Sequence 4, Application US/10024632  
; Publication No. US20020170093A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology LLC  
; APPLICANT: He, Steve S.  
; APPLICANT: Dotson, Stanton B.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION  
; FILE REFERENCE: 38-21(51837)B  
; CURRENT APPLICATION NUMBER: US/10/024,632  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: US 60/257,896  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 665  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-10-024-632-4

Query Match 49.0%; Score 1715; DB 13; Length 665;  
Best Local Similarity 54.4%; Pred. No. 2, 2e-151;  
Matches 379; Conservative 84; Mismatches 166; Indels 68; Gaps 20;

QY 1 MKAINSNNTDDGNHNMWLGFSLSPHM-----MKETSAATVPTTFYMPSPQSQSHSN 52  
DB 1 MKMENDDNADLNQNNWLGFSLSPOHNTGVSHSQPSAAEAVTSTSYHTHTAP--LSS 58  
QY 53 GGNVCYGV-GENGNFHSPLTVMPLKSDGSLCILEALKRSQTVVVPSTPKLEDPLGAGTM 111  
DB 53 YGFFYIGLEANNVGLYSALPIMPPLKSDGSLYGLTILRSQAQAWATSTPKLENFILGCEAM 118

QY 112 GT-HEY---GSHERGLSDSIYVNSQNAEAPNBDLSQPERQ-----QGMVSQTHPY 162  
DB 112 GT-HEY---GSHERGLSDSIYVNSQNAEAPNBDLSQPERQ-----QGMVSQTHPY 162  
QY 129 GTPHHYEGSATETWPLSLDSVYF--TQPRBDPNNQTVQNHVQHTSTNQOQOQOQELQAY 177  
DB 129 GTPHHYEGSATETWPLSLDSVYF--TQPRBDPNNQTVQNHVQHTSTNQOQOQOQELQAY 177  
QY 163 SGLACHGLYQAPLEETTKEHVSQSSLMQ-----KTEGLXNVAVPTRE--STHQ 213  
DB 163 SGLACHGLYQAPLEETTKEHVSQSSLMQ-----KTEGLXNVAVPTRE--STHQ 213  
QY 178 STLRHNDML-----EGSKQSTSDNNLHVQNMGGDDAVPVGLKSN--EVRNFQASHA 230  
DB 178 STLRHNDML-----EGSKQSTSDNNLHVQNMGGDDAVPVGLKSN--EVRNFQASHA 230  
QY 214 QVLEQNMCMGNENGV--SLGSVCGELOSLSMSPGSSCVT-----APSGTDSV 266  
DB 214 QVLEQNMCMGNENGV--SLGSVCGELOSLSMSPGSSCVT-----APSGTDSV 266  
QY 231 H--EGMVIYHVEENAGSGSIGMAYPDLOSLSMSPGSSSVTSSHRASPAVDSV 288  
DB 231 H--EGMVIYHVEENAGSGSIGMAYPDLOSLSMSPGSSSVTSSHRASPAVDSV 288  
QY 267 AYDAKRGHAKLGQKQPVHRKSIDTFGQRTSQRVTRHRTGRYEALHMDNSCKKEGT 326  
DB 267 AYDAKRGHAKLGQKQPVHRKSIDTFGQRTSQRVTRHRTGRYEALHMDNSCKKEGT 326  
QY 289 AMDTKRGPEKVDQKQVHRKSIDTFGQRTSQRVTRHRTGRYEALHMDNSCKKEG 348  
DB 289 AMDTKRGPEKVDQKQVHRKSIDTFGQRTSQRVTRHRTGRYEALHMDNSCKKEG 348  
QY 327 RKGROYLGGYDMEKARAYDLALAKTWGPSTHINFSTENYQVLEEMKMSRQEVYAH 386  
DB 327 RKGROYLGGYDMEKARAYDLALAKTWGPSTHINFSTENYQVLEEMKMSRQEVYAH 386  
QY 349 RKGROYLGGYDMEKARAYDLALAKTWGPSTHINFSTENYQVLEEMKMSRQEVYAH 408  
DB 349 RKGROYLGGYDMEKARAYDLALAKTWGPSTHINFSTENYQVLEEMKMSRQEVYAH 408  
QY 387 LRKSSGFSRGASIRGVTRHHQGRWQARIGRVANCKDLYLGTSTOEEAAEAYDAI 446  
DB 387 LRKSSGFSRGASIRGVTRHHQGRWQARIGRVANCKDLYLGTSTOEEAAEAYDAI 446  
QY 409 LRKSSGFSRGASIRGVTRHHQGRWQARIGRVANCKDLYLGTSTOEEAAEAYDAI 468  
DB 409 LRKSSGFSRGASIRGVTRHHQGRWQARIGRVANCKDLYLGTSTOEEAAEAYDAI 468  
QY 447 KRGANAVTNPDISRYDVERIMASSNLLAGELARXNDPBNKIDYKRSVYTVNNE 505  
DB 447 KRGANAVTNPDISRYDVERIMASSNLLAGELARXNDPBNKIDYKRSVYTVNNE 505  
QY 469 KRGANAVTNPDISRYDVERIMASSNLLAGELARXNDPBNKIDYKRSVYTVNNE 528  
DB 469 KRGANAVTNPDISRYDVERIMASSNLLAGELARXNDPBNKIDYKRSVYTVNNE 528  
QY 506 ETVOYQAGNNDNDEEMKQVLFNHPSCQOQANGSDOKIMQGNYSNRSFAMALQDL 565  
DB 506 ETVOYQAGNNDNDEEMKQVLFNHPSCQOQANGSDOKIMQGNYSNRSFAMALQDL 565  
QY 529 EALMHQKCESEND--QWIXVLY--QSSQLEQNPPTIE-----SDRTQSFVALDNVF 580  
DB 529 EALMHQKCESEND--QWIXVLY--QSSQLEQNPPTIE-----SDRTQSFVALDNVF 580  
QY 566 GIDVSGQHMLDESSKIGTFHFSNTSLVTSLSRSEASPEKRGPSLLIPMPMETKIV 625  
DB 566 GIDVSGQHMLDESSKIGTFHFSNTSLVTSLSRSEASPEKRGPSLLIPMPMETKIV 625  
QY 581 -----HQEVESSKARTHVSNPSLATSLSRSEASPEKRGPSLLIPMPMETKIV 631  
DB 581 -----HQEVESSKARTHVSNPSLATSLSRSEASPEKRGPSLLIPMPMETKIV 631  
QY 626 NPIGTSVSWLPSPVQMPSPALISHLPVFASWTD 662  
DB 626 NPIGTSVSWLPSPVQMPSPALISHLPVFASWTD 662  
QY 632 ATNPNVNSWSDPSFHLR-----PALTLPOMFVFAWTD 664  
DB 632 ATNPNVNSWSDPSFHLR-----PALTLPOMFVFAWTD 664

## RESULT 3

US-10-024-632-6  
; Sequence 6, Application US/10024632  
; Publication No. US20020170093A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology LLC  
; APPLICANT: He, Steve S.  
; APPLICANT: Dotson, Stanton B.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION  
; FILE REFERENCE: 38-21(51837)B  
; CURRENT APPLICATION NUMBER: US/10/024,632  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: US 60/257,896  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 642  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-024-632-6

Query Match 40.2%; Score 1408.5; DB 13; Length 642;  
Best Local Similarity 46.5%; Pred. No. 1, 1e-122;  
Matches 341; Conservative 64; Mismatches 149; Indels 179; Gaps 27;

QY 13 GNNHNMWLGFSLSPHMKEATSAATVPTT-----F 41  
DB 5 GGNSSNMWLGFSLSPHMKEATSAATVPTT-----F 41  
QY 42 YMSPSQSHL-----SNFMCYGVGEN--GNFHSPLTVMPLKSDGSLCILEALKRSQTV 92

```

Db      65 LFPBPAQWVAPSPGYVYGAYGDTSTAGVYSHLPVMPFIKSDSLCIMEG----- 117
Qy      93 VMTPTSSPKLEDLGCATWGTHERGHERLSDSYNSQNAEAQPNDDLISQPRROG 152
Db      118 -NMPSSPKLEDELGCNGSGH-----DPTTYSQGOEAD-----ASRAAYH 160
Qy      153 HNSVQTHPYVSGHAGLYQAPLEETTKETHVSDSSIMPOMTEGLKNWVAPTREFSTH 212
Db      161 HOLV---FVN-----YQPLTEAWLQEAAPAMEDM-----AAAKF--- 195
Qy      213 QQVLEQMGCMGNERNGVSLGVCCEL-QSISLMSPGSQ--SSCVT-APSGTSDVAD 269
Db      196 ---LVTSYGACYGNO-----EMPOPISLSPSPSSSCVSAAPQOHOQWAV 240
Qy      270 A-----KKRGAHLGQKQPVHRKSIDTFQORTSQYRGVTRHRM 307
Db      241 AAAAAADGGGNSNDGEGQVGRVKGKGTGKGQKQPVHRKSIDTFQORTSQYRGVTRHRM 300
Qy      308 TGRYEALHMDNSCKEGQTKRGQVYLGTYDMEKARAYDLAALKYNGPSTHINSIN 367
Db      301 TGRYEALHMDNSCKEGQTKRGQVYLGTYDMEKARAYDLAALKYNGPSTHINSIN 360
Qy      368 YQVLEEMKMSRQRYVAHLRRKSGSPSRGASITRGVTRHOGHGRQARIGRVAGNKDLY 427
Db      361 YRDEIEEMERTKQRYVAHLRRKSGSPSRGASITRGVTRHOGHGRQARIGRVAGNKDLY 420
Qy      428 LGTSTOEAAAYDAVAIKFRGANAVTNFDSRYVERIMASSNLIGELARRK--- 483
Db      421 LGTSTOEAAAYDAVAIKFRGANAVTNFDSRYVERIMASSNLIGELARRK--- 480
Qy      484 --DNDPRNKQIDYKSVVTSVNNETVQVQAGNNNNENDSEWKKVLFNHSSQOQA--- 537
Db      481 APDPVPIGRLE-----GATEBASAAATVGT-----TDMKVL--HSSQOQAAACT 523
Qy      538 NGNSDQKIMNCGYRNASAFSMALQDIDISVSGGQHNMLDESKI-GTHFSNTSLVT 596
Db      524 EATMDLQK-----GFMDPAS--ALHGIVGFVESAADEIDVPGKISGINFSSSLVT 577
Qy      597 SLSSRSRSPKRPKPSLIFPMPMETKIVNPIGTSVTS---WLPSPVQ---MRSPAI 649
Db      578 SLSSRSRSPKRPKPSLIFPMPMETKIVNPIGTSVTS---WLPSPVQ---MRSPAI 630
Qy      650 SLSHLPVFASWTD 662
Db      631 --AHLPVFAAWTD 641

```

## RESULT 4

```

US-10-024-632-9
Sequence 9, Application US/10024632
Publication No. US20020170093A1
GENERAL INFORMATION:
APPLICANT: Monsanto Technology LLC
APPLICANT: He, Steve S.
APPLICANT: Dotson, Stanton B.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION A
TITLE OF INVENTION: GROWTH AND USES THEREOF
FILE REFERENCE: 38-21(51837)B
CURRENT FILING DATE: 2001-12-19
CURRENT APPLICATION NUMBER: US/10/024,632
PRIOR APPLICATION NUMBER: US 60/257,896
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 669
TYPE: PRT
ORGANISM: Oryza sativa
US-10-024-632-9

```

Query Match 38.5%; Score 1349.5; DB 13; Length 669;  
Best Local Similarity 45.3%; Pred. No. 4e-117;

```

Matches 341; Conservative 66; Mismatches 142; Indels 203; Gaps 34;
Qy      13 GNNHNLGFSLSPKRKEAT-----SATVPTTFWSP 45
Db      18 GGVGAMLGFSLSFMAVATYACAGVDVGHNNHHVHOQOOGGGLFYNPAAVASFYGG 77
Qy      46 SQ-----SHLSNFMGCVGNGFNHSLPLVPLKSDSLCTLEALK--RSQTVMPFSS 99
Db      78 GHDAVNTSAAGGSYTGAG-----FSSNPLSDSLCIMEALRGDGEQOQVVSAS 129
Qy      100 PKLEDPLG-GATWGTHERGHERLSD--SIYNSQNAEAQPNDDLISQPRROGHN-- 154
Db      130 PKLEDPLGAGAM-----ALSLDNSAFYGGCH--HHQHPAD 165
Qy      155 --SVQTHPYVSG--LACHGLYQAPLEETTKETHVSDSSIMPOMTEGLKNWVAPTREF 209
Db      166 GAVAGGDHGHGGGFLQC-----AVITGAGAG----- 192
Qy      210 STHQVLEQGMNC-----GMGNERNGVSLGVC-----GELQSLISM- 248
Db      193 --HDAALVHDQSAVAAGMAAHGGGVDIAAADAADY-CAAGPIIPFGHLPLTSM 249
Qy      249 SPGSQSCVLT--APSGTSDVAVDA--KKRGA-KLGQKQPVHRKSIDTFQORTSQYRGV 302
Db      250 SAGSOSCVTVQAAAAGPYYAMDAVSKRGGADPAGQKQPVHRKSIDTFQORTSQYRGV 309
Qy      303 TRHWTGTYEALHMDNSCKEGQTKRGQVYLGTYDMEKARAYDLAALKYNGPSTHIN 362
Db      310 TRHWTGTYEALHMDNSCKEGQTKRGQVYLGTYDMEKARAYDLAALKYNGPSTHIN 369
Qy      363 FSIENYQVLEEMKMSRQRYVAHLRRKSGSPSRGASITRGVTRHOGHGRQARIGRVAG 422
Db      370 FLEDYQELIEEMKMSRQRYVAHLRRKSGSPSRGASITRGVTRHOGHGRQARIGRVAG 429
Qy      423 NKDYLIGF-----STQEEAAEYDAVAIKFRGANAVTNFDSRYVERIM 468
Db      430 NKDYLIGFIFASAFAPARRAHAGTQEEAAEYDAVAIKFRGANAVTNFDTIRYVDKIL 489
Qy      469 ASSNLIGELARRK-KDNDPRNKQIDYKSVVTSVNNETVQVQAGNNNNENDSEWKKVLF 527
Db      490 ESSTLLPGELARRKKGVDGGG-----AAVAADAAALVQAG-----NVLEMKAT 535
Qy      528 -----FNHPSQOQANGSDQKIMNCGYRNASAFSMALQDIDISVSGGQHNMLDES 581
Db      536 AALPAAARTEQOQHGQGHQ--HDDLPSDASV-LQDITVVDAAAGAPR----- 585
Qy      582 SKIGTHFSNTSLVTSLSSSRSRSPKRP-----SLFPMPMETKIVNPIGTSV 632
Db      586 ---APHMSMA---TSLGNSRQSPD-RGVGGGGGGVLAFLFAKPAASRLYSBV--PL 636
Qy      633 TSWL-DSPTVQNRPSPA-TSLSHLPVFASWTD 662
Db      637 NTWASPSFAVSSVPFAGVSIHLPFAAWTD 668

```

## RESULT 5

```

US-10-059-911-2
Sequence 2, Application US/10059911
Publication No. US20030159180A1
GENERAL INFORMATION:
APPLICANT: Fischer, Robert L.
APPLICANT: Mizukami, Yukiko
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Methods for Altering Organ Mass in Plants
CURRENT FILING DATE: 2002-06-17
CURRENT APPLICATION NUMBER: US/10/059,911
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 555
TYPE: PRT
ORGANISM: Arabidopsis thaliana

```

FEATURE:  
OTHER INFORMATION: Arabidopsis ANTIFUNGAL (ANT)  
US-10-059-911-2

Query Match 37.9%; Score 1326; DB 14; Length 555;  
Best Local Similarity 43.5%; Pred. No. 3,1e-115;  
Matches 314; Conservative 69; Mismatches 113; Indels 226; Gaps 22;

```

QY 1 MKRINSSNTDDGNNHNLGFSLSPH-MKM-----EATSAATVPTTFYM 43
DB 1 MKSFCDNDNNHNTNLTGFSLSNNMKGGRGREAIYSSSTSSAATSSSVPPQLV 60
QY 44 SPSGHLNFGMCYGVGENGFHSPPLYMPLKSDGSLCILALKRSGTQVMWPTSSPKLE 103
DB 61 G---DNTSNFGVCYGSNPGIYSHMSVPLRSDGSLCMEALNRSSSHNHODSSPKYE 117
QY 104 DFLGATWGTHEYSHERG--LSLDSIYNSQNAEAQPNRDLISQPF---RQGHMSVQ 157
DB 118 DFFG---THNNTSHKEMDLSDLSLFYNTTH---EPNTTNFQEFFSPQTRNH--- 166
QY 158 THPYSGLAHGLYQAPLEETTKETHVSDCSLMPONTGLKNWVAPTRFSTHQVLE 217
DB 167 -----EERT--RNYGND-----PSLTHG----- 182
QY 218 QQNNCGMNERNGVSLGSGGEL-QSLSLSPSGSSGCVTAPS----- 261
DB 183 -----GSFNVGV-----YGFQOQSLSLSPSGSSGCVTAPS----- 261
QY 262 -----GDSVAVDAKKRGH---AKLGOKQPVHRKSIDTGGRTSQRYGTR 304
DB 230 QOISEALVETSVGFETTMAAKKKRQGEDVAVVQKQIVHRKSIDTGGRTSQRYGTR 289
QY 305 HRWTRRYEALHMDNSCKEKGQTRKRGQVYLGQYDMEKARAAYDLALKYWGSPSTHIFS 364
DB 290 HRWTRRYEALHMDNSCKEKGQTRKRGQVYLGQYDMEKARAAYDLALKYWGSPSTHIFS 349
QY 365 IENYQVLEEMKMSQOEYVAHLRRKSSGFSRGSASTYRGVTRHHQGRQOARIGRVAGNK 424
DB 350 AENYQKEIEDMKMTRQOEYVAHLRRKSSGFSRGSASTYRGVTRHHQGRQOARIGRVAGNK 409
QY 425 DLYLGTFTQOEBAEAAYVAALKFRGANAVTFDISRYDVERIMASNTLAGELARKKD 484
DB 410 DLYLGTFTQOEBAEAAYVAALKFRGTNAVTFDISRYDVERIMASNTLAGELARKKD 467
QY 485 NDPRKQIDYNSKSVTVSNNEETVQVQAGNNNNENDSEMKVTLFNHPSQOOQANGSGSQ 544
DB 468 -----NNSIV--VRNTE-----DQ 479
QY 545 KIMNCGNRYNSAFSMALODLIGIDVSGQHMLDESSKIGTHFSNTSLVTSLSSSREA 604
DB 480 TALNA-----VEGGSNREV 494
QY 605 SPEKGPST--LFPMPMETKIY-NPIGNSVTSMLESPVQMRPSPALISLHPVFAST 661
DB 495 STPERLISFPAIPALPOVQVKYFGSNMGMSFWTSPNAELK-TVALTLQMPVFAAMA 553
QY 662 DT 663
DB 554 DS 555

```

RESULT 6  
US-10-059-911-21  
Sequence 21, Application US/10059911  
Publication No. US20030159180A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Mizukami, Yukiko  
TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
FILE REFERENCE: 023070-090730US  
CURRENT APPLICATION NUMBER: US/10/059,911  
CURRENT FILING DATE: 2002-06-17

NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 21  
LENGTH: 555  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: structurally  
US-10-059-911-21

Query Match 37.7%; Score 1319; DB 14; Length 555;  
Best Local Similarity 43.2%; Pred. No. 2.2e-114;  
Matches 312; Conservative 70; Mismatches 114; Indels 226; Gaps 22;

```

QY 1 MKRINSSNTDDGNNHNLGFSLSPH-MKM-----EATSAATVPTTFYM 43
DB 1 MKSFCDNDNNHNTNLTGFSLSNNMKGGRGREAIYSSSTSSAATSSSVPPQLV 60
QY 44 SPSGHLNFGMCYGVGENGFHSPPLYMPLKSDGSLCILALKRSGTQVMWPTSSPKLE 103
DB 61 G---DNTSNFGVCYGSNPGIYSHMSVPLRSDGSLCMEALNRSSSHNHODSSPKYE 117
QY 104 DFLGATWGTHEYSHERG--LSLDSIYNSQNAEAQPNRDLISQPF---RQGHMSVQ 157
DB 118 DFFG---THNNTSHKEMDLSDLSLFYNTTH---EPNTTNFQEFFSPQTRNH--- 166
QY 158 THPYSGLAHGLYQAPLEETTKETHVSDCSLMPONTGLKNWVAPTRFSTHQVLE 217
DB 167 -----EERT--RNYGND-----PSLTHG----- 182
QY 218 QQNNCGMNERNGVSLGSGGEL-QSLSLSPSGSSGCVTAPS----- 261
DB 183 -----GSFNVGV-----YGFQOQSLSLSPSGSSGCVTAPS----- 261
QY 262 -----GDSVAVDAKKRGH---AKLGOKQPVHRKSIDTGGRTSQRYGTR 304
DB 230 QOISEALVETSVGFETTMAAASSTRQGEDVAVVQKQIVHRKSIDTGGRTSQRYGTR 289
QY 305 HRWTRRYEALHMDNSCKEKGQTRKRGQVYLGQYDMEKARAAYDLALKYWGSPSTHIFS 364
DB 290 HRWTRRYEALHMDNSCKEKGQTRKRGQVYLGQYDMEKARAAYDLALKYWGSPSTHIFS 349
QY 365 IENYQVLEEMKMSQOEYVAHLRRKSSGFSRGSASTYRGVTRHHQGRQOARIGRVAGNK 424
DB 350 AENYQKEIEDMKMTRQOEYVAHLRRKSSGFSRGSASTYRGVTRHHQGRQOARIGRVAGNK 409
QY 425 DLYLGTFTQOEBAEAAYVAALKFRGANAVTFDISRYDVERIMASNTLAGELARKKD 484
DB 410 DLYLGTFTQOEBAEAAYVAALKFRGTNAVTFDISRYDVERIMASNTLAGELARKKD 467
QY 485 NDPRKQIDYNSKSVTVSNNEETVQVQAGNNNNENDSEMKVTLFNHPSQOOQANGSGSQ 544
DB 468 -----NNSIV--VRNTE-----DQ 479
QY 545 KIMNCGNRYNSAFSMALODLIGIDVSGQHMLDESSKIGTHFSNTSLVTSLSSSREA 604
DB 480 TALNA-----VEGGSNREV 494
QY 605 SPEKGPST--LFPMPMETKIY-NPIGNSVTSMLESPVQMRPSPALISLHPVFAST 661
DB 495 STPERLISFPAIPALPOVQVKYFGSNMGMSFWTSPNAELK-TVALTLQMPVFAAMA 553
QY 662 DT 663
DB 554 DS 555

```

RESULT 7  
US-10-059-911-22  
Sequence 22, Application US/10059911  
Publication No. US20030159180A1  
GENERAL INFORMATION:



APPLICANT: Fischer, Robert L.  
 APPLICANT: Mizukami, Yukiko  
 TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
 FILE REFERENCE: 023070-090730US  
 CURRENT APPLICATION NUMBER: US/10/059,911  
 CURRENT FILING DATE: 2002-06-17  
 NUMBER OF SEQ ID NOS: 42  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 22  
 LENGTH: 555  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: structurally  
 OTHER INFORMATION: altered ANT protein ANTmutII, temperature sensitive  
 OTHER INFORMATION: mutant  
 US-10-059-911-22

Query Match 37.5%; Score 1313; DB 14; Length 555;  
 Best Local Similarity 43.1%; Pred. No. 7.9e-114;  
 Matches 311; Conservative 70; Mismatches 115; Indels 226; Gaps 22;

1 MKRINESNNDDGNNHMLGFSLSPIH-MKM-----EATSAATVPTTFM 43  
 1 MKSPCDNDNNHSTNTNLGFSLSNNMKRGGRGREALYSSSTSAATSSSVPLV 60  
 44 SPSGSHSLNFGMCGVGENGFHSPFLTYMPLKSDGSLCILEALRSQTVVVPSPKLE 103  
 61 G---DNTSNFVYCISNNGIYSHMSWPLRSDGSLCILEALRSQTVVVPSPKLE 117  
 104 DFLGATVGTGHEGSHERG--LSLDSIYNSQNAEAPNDLSQPR---RQGGHNSVQ 157  
 118 DFFG---THNNTSHKAMDLSLDSLFYNTTH---EENTNTNFQEFFSPQTRNH--- 166  
 158 THPIYSLGACGLVAPLEETETKETHNSDCSILMPQTEGLKRWVAPTRFSTHQVLE 217  
 167 -----EELT---RNYGND-----PSLTHG----- 182  
 218 QQNNCGMNERNGVSLGSGGEL-QSLSLMSPSGSSCVTAPS----- 261  
 183 -----GSPNVG-----YGEFQSLSLMSPSGSSCVTAPS----- 229  
 262 -----GTDVAVDAAKRGH---AKLGQKQPVHRSKIDTRFGQTSQYRGVTR 304  
 230 QOISEALVETSVGFETTTMAAASSTRGQEDVVVGQKQIVHRSKIDTRFGQTSQYRGVTR 289  
 305 HMTGRVYEAHLMDNSCKEGOTRKGROYLLGQYDMEERAKARAYDLAALKTWGPSTHINS 364  
 230 HMTGRVYEAHLMDNSCKEGOTRKGROYLLGQYDMEERAKARAYDLAALKTWGPSTHINS 349  
 365 IENVQVLEEMKNSRQRYEYVAHLRKSGSGFSGASITRGVTRHQHGMQARIGRVAGNK 424  
 350 AENYQKEIEMKNSRQRYEYVAHLRKSGSGFSGASITRGVTRHQHGMQARIGRVAGNK 409  
 425 DLYLGTFSQGEAAAYDVAALIKFGANAVTNFDISRDVERIMASSNLAGEIARRKCD 484  
 410 DLYLGTFSQGEAAAYDVAALIKFGANAVTNFDISRDVERIMASSNLAGEIARRKCD 467  
 485 NDRNKDIDYNSKSVTVNNEETVQVQAGNNNNENDESEWKVLFNHPSCQOQOANGNSDQ 544  
 468 -----NNSIV---VANTE-----DQ 479  
 545 KIMNCGNYSASFMAQLDILIGDSVSGQHMLDESSKIGTHFSNTSLVTSLSRREA 604  
 480 TLANA-----VVEGSSNKEV 494  
 605 SPEKQPSL--LFPMPMETKIV-NPIGTSTWLPSTVQNPSPALISHLVYPAWST 661  
 495 STPERLSPPALFALPOVQKMFSGNMGNSPMTSNPAELK-TVALTLPOMPVFAMA 553  
 662 DT 663

DB 554 DS 555

RESULT 8  
 US-10-059-911-5  
 Sequence 5, Application US/10059911  
 Publication No. US20030159180A1  
 GENERAL INFORMATION:  
 APPLICANT: Fischer, Robert L.  
 APPLICANT: Mizukami, Yukiko  
 TITLE OF INVENTION: The Regents of the University of California  
 FILE REFERENCE: 023070-090730US  
 CURRENT APPLICATION NUMBER: US/10/059,911  
 CURRENT FILING DATE: 2002-06-17  
 NUMBER OF SEQ ID NOS: 42  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 5  
 LENGTH: 548  
 TYPE: PRT  
 ORGANISM: Brassica napus  
 FEATURE:  
 OTHER INFORMATION: Canola AINTGUMENTA (ANT)  
 US-10-059-911-5

Query Match 37.3%; Score 1307; DB 14; Length 548;  
 Best Local Similarity 43.6%; Pred. No. 2.8e-113;  
 Matches 309; Conservative 77; Mismatches 103; Indels 220; Gaps 25;

9 NTDDGNNHMLGFSLSPIH-----KMEATSAATVPTTFYWSPEQ---SHLSNFGMCY 57  
 6 DNDSDVTTLLGFSLSNNLKMGGGREALYSSSSSSSVANS--SVFQVYVGDSSNVGYCY 63  
 58 GVG-ENGAFHSPVLYMPLKSDGSLCILEALKS-----QTVVVPSPKLEDFLGCA 109  
 64 GSNLAREWYSQMSWVPLRSDGSLCILEALNRSQSHNNHHSQV---SSPQKEDF--- 115  
 110 TWGTHYEG-SHERG--LSLDSIYNSQNAEAPNDLSQPRQGGHNSVQTHPYSGLA 166  
 116 -FETHHNTSHKAMDLSLDSLFYNTTH---PNNNTNFQEF---FSPQTRNH--- 163  
 167 CHGLVAPLEETETKETHNSDCSILMPQTEGLKRWVAPTRFSTHQVLEQOQNGMGN 226  
 164 -----EELT---RNYGND-----PSLTHG----- 179  
 227 ERNGVSLGSGGEL-QSLSLMSPSGSSCVTA-----P 260  
 180 ---GGSFNVGVYGEFQSLSLMSPSGSSCVTAPS----- 236  
 261 SGTDSVAVDAAKRGH-HAKLGQKQPVHRSKIDTRFGQTSQYRGVTRHMTGRVYEAHLMDNS 319  
 237 TTTMAAALKKRGQGVVVVGQKQIVHRSKIDTRFGQTSQYRGVTRHMTGRVYEAHLMDNS 296  
 320 CKKEGOTRKGROYLLGQYDMEERAKARAYDLAALKTWGPSTHINSFINENYQVLEEMKNS 379  
 297 FKKEGSRGRROYLLGQYDMEERAKARAYDLAALKTWGPSTHINSFINENYQVLEEMKNS 356  
 380 ROEYVAHLRKSGSGFSGASITRGVTRHQHGMQARIGRVAGNKDLYLGTFSQGEAAE 439  
 357 ROEYVAHLRKSGSGFSGASITRGVTRHQHGMQARIGRVAGNKDLYLGTFSQGEAAE 416  
 440 AYDVAALIKFGANAVTNFDISRDVERIMASSNLAGEIARRKCDNDRNKDIDYNSKSV 499  
 417 AYDVAALIKFGANAVTNFDISRDVERIMASSNLAGEIARRKCDNDRNKDIDYNSKSV 465  
 500 TSYNNEE---TVQVQAGNNNNENDESEWKVLFNHPSCQOQOANGNSDQKIMNCGNYSASF 556  
 466 RNISDEEALITAVYNGSNKEV-----GSPFVLSF----- 496  
 557 FSNALDILIGDSVSGQHMLDESSKIGTHFSNTSLVTSLSRREASPEKQPSLFP 616  
 497 -----PTIFA 501

```

Qy      617  MPPMETKI--VNPICGTSVTSWLPSPFVQMRPSPALISLHPVEASMTDT 663
          :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      502  LPQVGPKMGANVVG-NMSSWTTNPADLK-TVSLITLPQMPVFPAAMADS 548

```

RESULT 9  
US-10-059-911-13

; Sequence 13, Application US/10059911

Publication No. US20030159180A1

GENERAL INFORMATION:

APPLICANT: Fischer, Robert L.

APPLICANT: Mizukami, Yukiko

APPLICANT: The Regents of the University of California

TITLE OF INVENTION: Methods for Altering Organ Mass in Plants

FILE REFERENCE: 023070-090730US

FILE REFERENCE: 023070-090/3005  
CURRENT APPLICATION NUMBER: US/10/059,911

CURRENT APPLICATION NUMBER: US/10/059,911  
CURRENT FILING DATE: 2003-06-17

CURRENT FILING DATE: 2002-06-17  
REVISED BY: CEO TO NO. 43

NUMBER OF SEQ ID NOS: 42

SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 13

; LENGTH: 452

TYPE: PRT

```

; LIFE: PRI
; ORGANISM: Artificial Sequence

```

```

; ORGANISM: Artificial sequence
FEATURE:
:

```

OTHER INFORMATION: Description of Artificial

OTHER INFORMATION

Query Match	35.2%	Score 1233.5	-DB 14	Length 452
Best Local Similarity	52.6%	Pred. No. 1.6e-106		
Matches 2/5	42	Mismatches 79	Indels 127	Gaps 15

```

QY      1 MRRINESNTDGGNNHNMJGFLSH-MM-----EAI$AVIvltvltm 43
Db      1 M$FCDNDNNH$NTNLLJGFL$LSNM$KMGGRGREALYSS$ST$SAAT$SS$VPPOLV 60
QY      4  $P$O$H$L$N$F$M$C$Y$G$V$G$N$G$N$H$P$L$Y$W$P$L$K$D$S$D$I$C$L$E$L$X$S$Q$V$W$V$P$S$P$K$E 103
Db      61  G---D$N$T$P$G$Y$C$S$N$P$N$G$I$Y$H$A$S$V$W$P$L$B$D$S$L$C$M$E$L$N$S$H$S$N$H$Q$D$S$P$K$E 117
QY      104  D$F$G$A$T$M$T$H$E$Y$G$S$H$E$R$G--L$D$S$I$Y$V$N$S$O$N$A$E$O$P$R$D$L$S$O$P$F---R$Q$G$H$M$S$V$Q 157
Db      118  D$F$G---T$H$A$N$T$S$H$K$E$A$M$D$L$S$D$L$F$Y$N$T$H---E$P$Y$T$T$N$F$O$E$F$S$P$O$T$R$N$H--- 166
QY      158  T$H$P$Y$S$G$I$A$C$H$G$Y$O$A$P$L$E$E$F$T$K$E$H$V$D$C$S$L$M$P$O$M$E$G$L$K$M$W$A$P$R$E$S$T$H$O$V$E 217
Db      167 -----E$E$T--R$N$Y$G$N$D---P$S$L$H$G----- 182
QY      218  Q$Q$M$G$M$G$N$E$R$N$G$V$S$L$G$S$V$G$C$E$L--Q$S$L$S$M$S$P$G$O$S$C$V$T$A$P$S----- 261
Db      183 -----G$F$N$Y$G$V-----Y$G$F$Q$S$L$B$M$S$P$G$O$S$C$I$G$H$N$H$Q$O$N$Q$N$H$Q$O$M$H 229
QY      262 -----G$T$D$V$A$V$D$A$K$K$G$H---A$K$G$O$K$O$V$N$H$K$S$I$T$F$O$R$S$O$Y$G$V$T$R 304
Db      230  Q$O$I$E$A$L$V$E$T$S$V$E$F$T$T$A$A$A$K$K$G$O$E$D$V$V$V$V$G$O$K$O$V$H$R$K$S$I$D$T$F$O$R$T$S$O$Y$G$V$T$R 289
QY      305  H$M$T$G$Y$E$A$H$M$D$N$S$C$K$E$G$O$T$K$G$R$O$V$Y$L$G$Y$D$M$E$K$A$R$A$D$L$A$L$K$Y$M$G$P$S$T$H$N$F$S 364
Db      290  H$M$T$G$Y$E$A$H$M$D$N$S$K$G$H$S$K$G$O$V$Y$L$G$Y$D$M$E$K$A$R$A$D$L$A$L$K$Y$M$G$P$S$T$H$N$F$S 349
QY      365  I$E$N$Y$O$V$L$E$M$K$M$S$R$O$E$Y$A$H$L$R$R$K$S$G$F$S$R$C$A$S$I$Y$G$V$T$R$H$O$H$G$E$W$O$A$R$I$G$Y$A$N$K 424
Db      350  A$E$N$Y$O$K$E$I$E$D$M$K$M$T$R$O$E$Y$A$H$L$R$R$K$S$G$F$S$R$C$A$S$I$Y$G$V$T$R$H$O$H$G$E$W$O$A$R$I$G$Y$A$N$K 409
QY      425  D$L$Y$G$F$S$T$O$E$A$A$E$Y$D$A$A$L$K$E$R$A$N$A$V$T$N$D$I$S$R$Y$C$E$R$I 467
Db      410  D$L$Y$G$F$G$T$O$E$A$A$E$Y$D$A$A$L$K$E$R$G$N$A$V$T$N$D$I$S$R$Y$C$E$R$I 452

```

RESULT 10  
US-10-059-911-9  
; Sequence 9, Application US/10059911  
; Publication No. US20030159180A1

```

// GENERAL INFORMATION:
// APPLICANT: Fischer, Robert L.
// APPLICANT: Mizukami, Yukiko
// APPLICANT: The Regents of the University of California
// TITLE OF INVENTION: Methods for Altering Organ Mass in Plants
// FILE REFERENCE: 023070-090730US
// CURRENT APPLICATION NUMBER: US/10/059,911
// CURRENT FILING DATE: 2002-06-17
// NUMBER OF SEQ ID NOS: 42
// SOFTWARE: PatentIn Ver. 2.1
// SEQ ID NO 9
// LENGTH: 473
// TYPE: PRT
// ORGANISM: Artificial Sequence
// FEATURE:
// OTHER INFORMATION: Description of Artificial Sequence:structurally
// /S-10-059-911-9

```

Query Match	35.0%	Score 1224	DB 14	Length 473
Best Local Similarity	44.9%	Pred. No. 1.3e-105		
Matches 286: Conservative	56	Mismatches	89	Indels 206
				Gaps 19

[illegible]

```

US-10-024-632-11
; Sequence 11, Application US/10024632
; Publication No. US20020170093A1
; GENERAL INFORMATION:

```

/ APPLICANT: Monsanto Technology LLC  
 / APPLICANT: He, Steve S.  
 / APPLICANT: Dotson, Stanton B.  
 / TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION  
 / TITLE OF INVENTION: GROWTH AND USES THEREOF  
 / FILE REFERENCE: 38-21(5183)B  
 / CURRENT APPLICATION NUMBER: US/10/024,632  
 / CURRENT FILING DATE: 2001-12-19  
 / PRIOR APPLICATION NUMBER: US 60/257,896  
 / PRIOR FILING DATE: 2000-12-21  
 / NUMBER OF SEQ ID NOS: 33  
 / SOFTWARE: PatentIn version 3.1  
 / SEQ ID NO 11  
 / LENGTH: 585  
 / TYPE: PRT  
 / ORGANISM: Gossypium hirsutum  
 / US-10-024-632-11

Query Match 33.7%; Score 1180.5; DB 13; Length 585;  
 Best Local Similarity 41.7%; Pred. No. 2,2e-101; Indels 163; Gaps 25;  
 Matches 250; Conservative 79; Mismatches 164;

QY 17 NMWGLSLSPHMKKATSAATVPTTFYMSPOSLSLNFQMCYGVGENGNFHSPLTWPLKS 76  
 DB 3 NMWGLSLTPDLRIDES-----FGREDHGFPSS---VMDLRS 35  
 QY 77 DGSICLLEALKRSQT-----QVWVPTSSPKLEDFLG---GATMGTHEY-GS 118  
 DB 36 DGSICLVDPFRSSIAADBDREYENGISATNDEGQPLDPLGCSNPSQETKAYCGT 95  
 QY 119 HER-----GLSLDSIYNSQNAAPN-----BDLSQP--FRQGHMS 155  
 DB 96 HENONTVPSPRTINVAAPNYSSSGDAEAENFTNPSFPIOTRYNNYENPQTLMAQHS 155  
 QY 156 VCHPEYSGLAGHGYQAPLEETTKETHVSDCSSLMQMEGLKNWAPRREHSTHQV 215  
 DB 156 QCCDPFNHNGSRGVAHVDFESATS-----VSGRKSMLRQT--PPGSK-- 197  
 QY 216 LEOQNNCGNENRNGVSLGSGELQSLSLMSBPSQSCVTAAPSQTSVYA---VDAK 271  
 DB 198 -----ASGNETNNNF-----NFOALSLTMSB-----TSRNGFPALAPLEVVDNR 236  
 QY 272 KSGHAKLGOKOVYHRKSIDTFGQRTSOYRGVTRHRTGRYEAHLMDSCKKESGTRKRG 331  
 DB 237 KRPVGNLTRRESVPRKSIDTFGQRTSOYRGVTRHRTGRYEAHLMDSCKKESGTRKRG 236  
 QY 332 VYLGVDYMEKARAYDLAALKYWGSPSTHINFSIENYQVLEEMKNMSROEYVAHLRR 391  
 DB 297 VYLGVDYMEKAKAYDLAALKYWGSPSTHINFSIENYQVLEEMKNMTRCEYVAHLRR 356  
 QY 392 SSGSRASLYRGVTRHGHGRQWARIIGVAGKXDLYLGTFTSOREAAEAYVAALIRGA 451  
 DB 357 SSGSRASLYRGVTRHGHGRQWARIIGVAGKXDLYLGTFTSOREAAEAYVAALIRGT 416  
 QY 452 NAYTNFDIRYVERIMASSNLLAGELARRKXNDPRNKDIDYKSVVTSVYNNETVQV 511  
 DB 417 SAVTNFDIRYVERIKRISSTLIGELAKRSPKDTASIAPEVDN-SCASSASQPLAIP 475  
 QY 512 AGNNNNE-NDSEKAYTLFPHSPQOO-QANGNSDDKINNCANYNNSAFVVALDLDLGI-- 567  
 DB 476 SGEASBELADWYTA---NSDEQOQHSTNTNNDASLANSSS-RNSNPSQSPKSTGLAS 531  
 QY 568 DVSQSGCHNMLDESSKIGTFPSNTSLVTSJSSSEASPEKRGPSLLFPMPEMETKLVNP 627  
 DB 532 DKRGID---GYSYHSG-YFSLKSGSKYEDGNSSEIDNSNENR-----567  
 QY 628 IGTSVTSMLPSPVQWRPSPALSLSH-LPVFASMTD 662  
 DB 568 LGN-----LGLVHKIIPMFALMNE 585

RESULT 12  
 US-10-059-911-15

/ Sequence 15, Application US/10059911  
 / Publication No. US20030159180A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Fischer, Robert L.  
 / APPLICANT: Mizukami, Yukiko  
 / APPLICANT: The Regents of the University of California  
 / TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
 / FILE REFERENCE: 023070-090730US  
 / CURRENT APPLICATION NUMBER: US/10/059,911  
 / CURRENT FILING DATE: 2002-06-17  
 / NUMBER OF SEQ ID NOS: 42  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 15  
 / LENGTH: 370  
 / TYPE: PRT  
 / ORGANISM: Artificial Sequence  
 / FEATURE:  
 / OTHER INFORMATION: Description of Artificial Sequence: structurally  
 / OTHER INFORMATION: altered ANT protein ANTDMCI  
 / US-10-059-911-15

Query Match 32.3%; Score 1129.5; DB 14; Length 370;  
 Best Local Similarity 56.4%; Pred. No. 6.5e-97;  
 Matches 247; Conservative 29; Mismatches 55; Indels 107; Gaps 12;

QY 69 LTVMPKSDGSLCILEALKRSQIQVWVPTSSPKLEDFLGATMGTHEYSGHERG--LSLD 126  
 DB 1 MSVMPLRSDGSLCILEALKRSHSSNNHODSPKVEDFG-----THNNNTSHKRAMLSLD 56  
 QY 127 SIYNSQNAEAPNREDLSQPF---RQGHMSVOTHPYVSGLAGHGYQAPLEETTKE 182  
 DB 57 SLFVNTTH---EPNTTNFOCFEFSFPQTRNH-----EEET--R 89  
 QY 183 TVSPDSSLMQMEGLKNWAPRREHSTHQVLEQNNCGMGNERNVSLGSGCEL- 241  
 DB 90 NYGND-----PSLTHG-----GSFNVGV-----YGEFQ 112  
 QY 242 QSLSLSPSGSQSCVTAAPS-----GTDSVAVDARK 272  
 DB 113 QSLSLSPSGSQSCITTSHHHCNONONHOSQNHQISALVETSVGFETTTMAAKKK 172  
 QY 273 RGH---ATLGOKOVYHRKSIDTFGQRTSOYRGVTRHRTGRYEAHLMDSCKKESGTRKRG 329  
 DB 173 RQGEVYVYVGGQKQYHRKSIDTFGQRTSOYRGVTRHRTGRYEAHLMDSCKKESGTRKRG 232  
 QY 330 RQVYLGVDYMEKARAYDLAALKYWGSPSTHINFSIENYQVLEEMKNMSROEYVAHLRR 389  
 DB 233 RQVYLGVDYMEKARAYDLAALKYWGSPSTHINFSIENYQVLEEMKNMTRCEYVAHLRR 292  
 QY 390 KSSGSRASLYRGVTRHGHGRQWARIIGVAGKXDLYLGTFTSOREAAEAYVAALIKFR 449  
 DB 293 KSSGSRASLYRGVTRHGHGRQWARIIGVAGKXDLYLGTFTSOREAAEAYVAALIKFR 352  
 QY 450 GNAVTNFDISRYDVERI 467  
 DB 353 GTNAVTNFDITRYDVDR 370

RESULT 13  
 US-10-059-911-10  
 / Sequence 10, Application US/10059911  
 / Publication No. US20030159180A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Fischer, Robert L.  
 / APPLICANT: Mizukami, Yukiko  
 / APPLICANT: The Regents of the University of California  
 / TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
 / FILE REFERENCE: 023070-090730US  
 / CURRENT APPLICATION NUMBER: US/10/059,911  
 / CURRENT FILING DATE: 2002-06-17  
 / NUMBER OF SEQ ID NOS: 42  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 10

LENGTH: 308  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: structurally  
OTHER INFORMATION: altered ANT protein ANTND2  
US-10-059-911-10

Query Match 29.8%; Score 1042; DB 14; Length 308;  
Best Local Similarity 54.0%; Pred. No. 7,7e-89;  
Matches 218; Conservative 36; Mismatches 48; Indels 102; Gaps 8;

QY 266 VAVDAKRGH---AKLGKQKPYHRSIDTFQGRISQYRGVTRHRTGTYEAHLMDNSCKK 322  
DB 1 MAIAKKKQDEVVVVGQKQIYHRSIDTFQGRISQYRGVTRHRTGTYEAHLMDNSCKK 60  
QY 323 ECGTRKRGROYIGGIDMEKARAYDLAKTKWGSTINFSIENYQVQLEEMKMGROE 382  
DB 61 EGHRSRGROYIGGIDMEKARAYDLAKTKWGSTINFSIENYQVQLEEMKMGROE 120  
QY 383 YVAHLRRKSGSGFSRGASIRGVTTRHOGHGRWQARIGVAGNDLYLGTFFQEEAAEAYD 442  
DB 121 YVAHLRRKSGSGFSRGASIRGVTTRHOGHGRWQARIGVAGNDLYLGTFFQEEAAEAYD 180  
QY 443 VAAIKRGANAVTNPDISRIYVERIMASNLLAGELARRKNDPRNDIDYNSVTSV 502  
DB 181 VAAIKRGANAVTNPDITRYDYDRIMSSNTLLSGELARRN---NNSIV--V 226  
QY 503 NNEETVQVQAGNNNNNENDESEWKKVLFNHPSSQOQAGNSDQKIMCGYNSAFSVALQ 562  
DB 227 RNE-----DGTALNA----- 237  
QY 563 DLIGDSVSGQHMLDESSKIGTHFSNTSLVTSLSGSRKSPKRGSL--LFPMPM 620  
DB 238 -----VVEGGSNKEVSTPERLLSPFAIFALPOV 265  
QY 621 ETKIV-NPIGTSVSWLPSPTVQMRPSPAISLSHLFPASWTDT 663  
DB 266 NQKMFSGNMGANSPWTSNPNELK-TVALTLPOMFVFAAMDS 308

RESULT 14  
US-10-059-911-11  
Sequence 11, Application US/10059911  
Publication No. US20030159180A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Mizukami, Yukiko  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
FILE REFERENCE: 023070-090730US  
CURRENT APPLICATION NUMBER: US/10/059,911  
CURRENT FILING DATE: 2002-06-17  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 11  
LENGTH: 277  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: structurally  
OTHER INFORMATION: altered ANT protein ANTND3  
US-10-059-911-11

Query Match 27.3%; Score 955.5; DB 14; Length 277;  
Best Local Similarity 53.2%; Pred. No. 8,3e-81;  
Matches 198; Conservative 34; Mismatches 41; Indels 99; Gaps 7;

QY 295 RTSQYRGVTRHRTGTYEAHLMDNSCKKGGQTRKRGROYIGGIDMEKARAYDLAKTK 354  
DB 2 RTSQYRGVTRHRTGTYEAHLMDNSCKKGGQTRKRGROYIGGIDMEKARAYDLAKTK 61  
QY 355 WGPSTHINFSIENYQVQLEEMKMGROEYVAHLRRKSGSGFSRGASIRGVTTRHOGHGRWQ 414

DB 62 WGPSTHINFSIENYQVQLEEMKMGROEYVAHLRRKSGSGFSRGASIRGVTTRHOGHGRWQ 121  
QY 415 ARIGRYAGNKDLYLGTFFQEEAAEAYVAIAKFRGANAVTNFDISRYDVERIMASNL 474  
DB 122 ARIGRYAGNKDLYLGTFFQEEAAEAYVAIAKFRGANAVTNFDITRYDVERIMASNL 181  
QY 475 AGEIARRKNDPRNKDIDYNSKVTSVNNEETVQVQAGNNNNNENDESEWKKVLFNHPSSQ 534  
DB 182 SGELARRN-----NNSIV--VANTE----- 199  
QY 535 QAGNSGSDQKIMCGYNSAFSVALQDLIGDSVSGQHMLDESSKIGTHFSNTSL 594  
DB 200 -----DGTALNA----- 206  
QY 595 VTSLSGSRKSPKRGSL--LFPMPMETKIV-NPIGTSVSWLPSPTVQMRPSPAISL 651  
DB 207 VVEGGSNKEVSTPERLLSPFAIFALPOVQKFGSNMGANSPWTSNPNELK-TVALTL 265  
QY 652 SHLPVPSWTDT 663  
DB 266 PQMPVFAAMDS 277

RESULT 15  
US-10-374-780A-334  
Sequence 334, Application US/10374780A  
Publication No. US20040019927A1  
GENERAL INFORMATION:  
APPLICANT: Sherman, Bradley K  
APPLICANT: Riechman, Jose Luis  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Heard, Jacqueline E  
APPLICANT: Haake, Volker  
APPLICANT: Creelman, Robert A  
APPLICANT: Raccilite, Oliver  
APPLICANT: Adam, Luc J  
APPLICANT: Reuber, T. Lynne  
APPLICANT: Keddie, James  
APPLICANT: Brown, Pierre E  
APPLICANT: Piggin, Marsha L  
APPLICANT: Dubell III, Arnold T  
APPLICANT: Phineda, Omaira  
APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS  
FILE REFERENCE: MBI-0047 CIP  
CURRENT APPLICATION NUMBER: US/10/374,780A  
CURRENT FILING DATE: 2003-02-25  
PRIOR APPLICATION NUMBER: 09/837,944  
PRIOR FILING DATE: 2001-04-18  
PRIOR APPLICATION NUMBER: 60/310,847  
PRIOR FILING DATE: 2001-08-09  
PRIOR APPLICATION NUMBER: 09/934,455  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/336,049  
PRIOR FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/338,692  
PRIOR FILING DATE: 2001-12-11  
PRIOR APPLICATION NUMBER: 10/171,468  
PRIOR FILING DATE: 2002-06-14  
PRIOR APPLICATION NUMBER: 10/225,066  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 10/225,067  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 10/225,068  
PRIOR FILING DATE: 2002-08-09  
NUMBER OF SEQ ID NOS: 2906  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 334  
LENGTH: 574  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:

OTHER INFORMATION: G1793 (conserved domain in AA coordinates:179-255, 281-349)  
US-10-374-780A-334

Query Match 27.1% Score 948; DB 15; Length 574;  
Best Local Similarity 35.9%; Pred No. 1,4e-79;  
Matches 257; Conservative 86; Mismatches 164; Indels 208; Gaps 26;

```
QY 14 NNNNTLIGPSLSPHMKMEATSAATVPTTFYMSPSQSHLSN-----FGMCYGVGENGNFHS 67
DB 2 NSNNMLGPPLSPN-----NSSLPHEVYVLGLVSDHMDNPFQTEMNNINPHGGGD--- 52
QY 68 PLTVMPKSDGLICILBAKRSQTVMPVTSSEPKLEDPLGATMGTHGYSGHERGLSLDS 127
DB 53 -----EGG-----EVPKADPLG--VSKPDENSHLVAYNDS 83
QY 128 IYNSQNAEAQPNRDLISQPFROGQMSVQTHPYYSGLACHGLYQAPLEETKETHVSD 187
DB 84 DYEFHTNS-LMP-----SYQSN-----DVVVA 105
QY 188 CSSLMPQMTGELKNWVAPTRPSTHOVLEQQNCGMGNERNNGVSLGSGGELQSLSLG 247
DB 106 CDSNTNPN-----SYHELQESNPN-----LQSLTSL 132
QY 248 MSPSGSSCVTAPSGTDSVAVDAKKRGAHKLQKQKQVHRKSIDTFGQRTSQYRGVTRRW 307
DB 133 MGTLAGNNVVDKASPSSETTGDNASGALAVETATP--RRALDTFGORTSIYRGVTRRW 190
QY 308 TGRYEALHMDNSCKEKGOTRKGRQVYLLGQYDMEEKARAYDIALKTYWGPSTHINFSIEN 367
DB 191 TGRYEALHMDNSCKEKGOTRKGRQVYLLGQYDMEEKARAYDIALKTYWGPSTHINFSIEN 250
QY 368 YQVQLEEMKNGRQEVVAHLRRKSSGFSGASIRGVTRHHQGRWQARIQVAGNKDLY 427
DB 251 YEKVEYEMKMTROEFVAIRKSSGFSGASIRGVTRHHQGRWQARIQVAGNKDLY 310
QY 428 LGTFSTQEEAEAYDVAATKFRGANAVTNFDISRVDVERIMASSNL-LAGELARKKD-- 484
DB 311 LGTFSTQEEAEAYDVAATKFRGANAVTNFDISRVDVERIMASSNL-LAGELARKKD-- 370
QY 485 --NDPNKCID-----YKNSV-----VTSVNNEET 507
DB 371 ALESSRKREAEEMIALGSSFOYGGSSSTGSGTSLRLQPYPLSLIQPLEPFLSLQVND- 429
QY 508 VOYQAGNNNNENDSEMKWVLFNHPGQ-----QQANGNGSDPKIMNGNY 552
DB 430 --ISHNNNNADDS---SFNHSYIQTQLHQQCTNNYLQOQSSON--SQCLYNAYLH 481
QY 553 RNSAFSMLQDLIGIDSYGSGQHNNLDESSKIGTHFSNTSLVTSL-----SSSRKSPPE 607
DB 482 SNPAL--LHGLVSTSIYDNNNNN---GSSGSGY--NTAFLGNHIGIGISSSTVGSTE 532
QY 608 KRGPILLFPMPEMKIYVPIGSTVSWLPSPTVQMRSPALISHLPVFASTWD 662
DB 533 E-----PPTYKTIDYDMPSSDGTGCGYSGWTSSEVO-GSNPG-----GVFTYMANE 574
```

Search completed: March 9, 2004, 10:52:46  
Job time: 40 secs